

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 09:22:17 ; Search time 0.731707 Seconds
(without alignments)
482.536 Million cell updates/sec

Title: US-09-890-752a-20
Perfect score: 61
Sequence: 1 PLSIFSRIGDP 12

Scoring table: BLOSUM62
Gap 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTCUTS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61	100.0	55	4	US-09-361-707-97
2	61	100.0	174	3	US-08-480-173A-44
3	61	100.0	174	3	US-08-484-408A-44
4	61	100.0	281	1	US-08-105-483-214
5	61	100.0	281	1	US-08-709-209-214
6	61	100.0	281	1	US-08-458-101-214
7	61	100.0	389	1	US-08-105-483-216
8	61	100.0	389	1	US-08-105-483-216
9	61	100.0	389	1	US-08-105-483-216
10	61	100.0	389	1	US-08-709-209-216
11	61	100.0	389	1	US-08-458-101-216
12	61	100.0	389	1	US-08-458-101-216
13	61	100.0	389	1	US-08-458-101-216
14	61	100.0	389	3	US-08-486-099-106
15	61	100.0	389	3	US-08-360-107A-116
16	61	100.0	389	3	US-08-484-223B-106
17	61	100.0	389	3	US-08-919-597-106
18	61	100.0	389	3	US-08-475-668A-106
19	61	100.0	389	3	US-08-485-551A-106
20	61	100.0	389	4	US-08-471-813A-106
21	61	100.0	389	4	US-08-485-264A-106
22	61	100.0	389	4	US-08-474-349A-106
23	59	96.7	55	4	US-09-361-707-93
24	59	96.7	55	4	US-09-361-707-95
25	59	96.7	55	4	US-09-361-707-96
26	59	96.7	174	2	US-08-683-262B-55
27	59	96.7	174	3	US-08-480-173A-45
				3	US-08-484-408A-45

28	59	96.7	174	4	US-09-361-707-55	Sequence 55, App1
29	59	96.7	174	6	5204096-1	Patent No. 5204096
30	59	96.7	174	6	5204096-2	Patent No. 5204096
31	54	88.5	14	6	5164485-4	Patent No. 5164485
32	54	88.5	55	4	US-09-361-707-84	Sequence 84, App1
33	54	88.5	55	4	US-09-361-707-85	Sequence 85, App1
34	54	88.5	55	4	US-09-361-707-87	Sequence 87, App1
35	54	88.5	55	4	US-09-361-707-89	Sequence 89, App1
36	54	88.5	55	4	US-09-361-707-90	Sequence 90, App1
37	54	88.5	170	2	US-08-683-262B-49	Sequence 49, App1
38	54	88.5	170	2	US-08-683-262B-46	Sequence 46, App1
39	54	88.5	174	2	US-08-683-262B-48	Sequence 48, App1
40	54	88.5	174	2	US-08-683-262B-48	Sequence 48, App1
41	54	88.5	174	2	US-08-683-262B-56	Sequence 56, App1
42	54	88.5	174	3	US-08-480-173A-48	Sequence 48, App1
43	54	88.5	174	3	US-08-484-408A-48	Sequence 48, App1
44	54	88.5	174	4	US-09-361-707-46	Sequence 46, App1
45	54	88.5	174	4	US-09-361-707-48	Sequence 48, App1

ALIGNMENTS

RESULT 1
US-09-361-707-97
Sequence 97, Application US/09361707
Patent No. 6258937

GENERAL INFORMATION:

APPLICANT: Tong, Shuping
L.I, Jisu
Wanda, Jack R.
TITLE OF INVENTION: HEPA DNA VIRUS RECEPTOR
CORRESPONDENCE ADDRESSES:
ADDRESSES: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/361,707
FILING DATE: 27-Jul-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/683,262
FILING DATE: 18-Jul-1996
ATTORNEY/AGENT INFORMATION:
NAME: Creason, Gary L.
REGISTRATION NUMBER: 34,310
REFERENCE/DOCKET NUMBER: 00786/287003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
INFORMATION FOR SEQ ID NO: 97:
SEQUENCE CHARACTERISTICS:
LENGTH: 55 amino acids
TOPOLOGY: linear

US-09-361-707-97
SEQUENCE DESCRIPTION: SEQ ID NO: 97:

Query Match 100.0%; Score 61; DB 4; Length 55;
Best Local Similarity 100.0%; Pred. No. 9.2e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PLSIFSRIGDP 12
DB 41 PLSIFSRIGDP 52

RESULT 2
US-08-480-173A-44
Sequence 44, Application US/08480173A
Patent No. 6072049
GENERAL INFORMATION:
APPLICANT: Thoma, Hans A
TITLE OF INVENTION: HEPATITIS B SURFACE ANTIGEN VACCINE
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Popovich & Wiles, P.A.
STREET: 80 S. 8th Street, Suite 1902
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480.173A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Popovich, Thomas E
REGISTRATION NUMBER: 30,099
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-334-8994
TELEFAX: 612-334-8994
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 174 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-480-173A-44

Query Match 100.0%; Score 61; DB 3; Length 174;
Best Local Similarity 100.0%; Pred. No. 0.00034;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFRIGDP 12
DB 160 PLSISFRIGDP 171

RESULT 3
US-08-484-408A-44
Sequence 44, Application US/08484408A
Patent No. 6117653
GENERAL INFORMATION:
APPLICANT: Thoma, Hans A
TITLE OF INVENTION: HEPATITIS B SURFACE ANTIGEN VACCINE
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Popovich & Wiles, P.A.
STREET: 80 S. 8th Street, Suite 1902
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484.408A

FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Popovich, Thomas E
REGISTRATION NUMBER: 30,099
REFERENCE/DOCKET NUMBER: MED1003USD4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-334-8994
TELEFAX: 612-334-8994
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 174 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-484-408A-44

Query Match 100.0%; Score 61; DB 3; Length 174;
Best Local Similarity 100.0%; Pred. No. 0.00034;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFRIGDP 12
DB 160 PLSISFRIGDP 171

RESULT 4
US-08-105-483-214
Sequence 214, Application US/08105483
Patent No. 5494807
GENERAL INFORMATION:
APPLICANT: Paolletti, Enzo
TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE
NUMBER OF SEQUENCES: 462
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford
ADDRESSEE: c/o William S. Frommer
STREET: 530 Fifth Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/105.483
FILING DATE: 12-AUG-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/847,951
FILING DATE: 06-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2400
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 214:
SEQUENCE CHARACTERISTICS:
LENGTH: 281 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-105-483-214

Query Match 100.0%; Score 61; DB 1; Length 281;
Best Local Similarity 100.0%; Pred. No. 0.0006;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSIFSRIGDP 12
Db 41 PLSIFSRIGDP 52

RESULT 5

US-08-709-209-214
Sequence 214, Application US/08709209
Patent No. 5762938
GENERAL INFORMATION:
APPLICANT: Paolietti, Enzo
TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE
NUMBER OF SEQUENCES: 462
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford
STREET: 530 Fifth Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08709,209
FILING DATE: 21-AUG-1996
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/105,483
FILING DATE: 12-AUG-1993
APPLICATION NUMBER: US 07/847,951
FILING DATE: 06-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2400
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 214:
SEQUENCE CHARACTERISTICS:
LENGTH: 281 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-709-209-214

Query Match Best Local Similarity 100.0%; Score 61; DB 1; Length 281;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSIFSRIGDP 12
Db 41 PLSIFSRIGDP 52

RESULT 6

US-08-458-101-214
Sequence 214, Application US/08458101
Patent No. 5766599
GENERAL INFORMATION:
APPLICANT: Paolietti, Enzo
TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE
NUMBER OF SEQUENCES: 462
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford
STREET: 530 Fifth Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

APPLICANT: de Taisne, Charles
APPLICANT: Limbach, Keith J.
APPLICANT: Johnson, Gerard P.
APPLICANT: Pincus, Steven E.
APPLICANT: Cox, William I.
APPLICANT: Audonnet, Jean-Christophe Francis
TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE
NUMBER OF SEQUENCES: 467
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford
STREET: 530 Fifth Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,101
FILING DATE: 01-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2740
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 214:
SEQUENCE CHARACTERISTICS:
LENGTH: 281 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-458-101-214

Query Match Best Local Similarity 100.0%; Score 61; DB 1; Length 281;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSIFSRIGDP 12
Db 41 PLSIFSRIGDP 52

RESULT 7

US-08-105-483-216
Sequence 216, Application US/08105483
Patent No. 5494807
GENERAL INFORMATION:
APPLICANT: Paolietti, Enzo
TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE
NUMBER OF SEQUENCES: 462
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford
STREET: 530 Fifth Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/105,483
FILING DATE: 12-AUG-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/847,951
FILING DATE: 06-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2400
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 216:
SEQUENCE CHARACTERISTICS:
LENGTH: 389 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-105-483-216

Query Match 100.0%; Score 61; DB 1; Length 389;
Best Local Similarity 100.0%; Pred. No. 0.00086;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSIFSRIGDP 12
DB 149 PLSIFSRIGDP 160

RESULT 8
US-08-105-483-219
Sequence 219, Application US/08105483
Patent No. 5494807
GENERAL INFORMATION:
APPLICANT: Paoletti, Enzo
TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE
TITLE OF INVENTION: STRAIN
NUMBER OF SEQUENCES: 462
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford
ADDRESSEE: c/o William S. Frommer
STREET: 530 Fifth Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/105,483
FILING DATE: 12-AUG-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/847,951
FILING DATE: 06-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2400
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-0712
TELEFAX: (212) 840-3333
INFORMATION FOR SEQ ID NO: 219:
SEQUENCE CHARACTERISTICS:
LENGTH: 389 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-105-483-219

Query Match 100.0%; Score 61; DB 1; Length 389;
Best Local Similarity 100.0%; Pred. No. 0.00086;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSIFSRIGDP 12
DB 149 PLSIFSRIGDP 160

RESULT 9
US-08-709-209-216
Sequence 216, Application US/08709209
Patent No. 5762938
GENERAL INFORMATION:
APPLICANT: Paoletti, Enzo
TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE
TITLE OF INVENTION: STRAIN
NUMBER OF SEQUENCES: 462
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford
ADDRESSEE: c/o William S. Frommer
STREET: 530 Fifth Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/709,209
FILING DATE: 21-AUG-1996
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/105,483
FILING DATE: 12-AUG-1993
APPLICATION NUMBER: US 07/847,951
FILING DATE: 06-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2400
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 216:
SEQUENCE CHARACTERISTICS:
LENGTH: 389 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-709-209-216

Query Match 100.0%; Score 61; DB 1; Length 389;
Best Local Similarity 100.0%; Pred. No. 0.00086;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSIFSRIGDP 12
DB 149 PLSIFSRIGDP 160

RESULT 10
US-08-709-209-219
Sequence 219, Application US/08709209
Patent No. 5762938
GENERAL INFORMATION:
APPLICANT: Paoletti, Enzo
TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE

TITLE OF INVENTION: STRAIN
NUMBER OF SEQUENCES: 462
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford
ADDRESSER: c/o William S. Frommer
STREET: 530 Fifth Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/709,209
FILING DATE: 21-AUG-1996
CLASSIFICATION: 424
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/105,483
FILING DATE: 12-AUG-1993
APPLICATION NUMBER: US 07/847,951
FILING DATE: 06-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2400
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 219:
SEQUENCE CHARACTERISTICS:
LENGTH: 389 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-709-209-219

Query Match 100.0%; Score 61; DB 1; Length 389;
Best Local Similarity 100.0%; Pred. No. 0.00086;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PLSIFSRIGDP 12
Db 149 PLSIFSRIGDP 160

RESULT 11
US-08-458-101-216
Sequence 216, Application US/08458101
Patent No. 5766599
GENERAL INFORMATION:
APPLICANT: Paoletti, Enzo
APPLICANT: Perkins, Marion E.
APPLICANT: Taylor, Jill
APPLICANT: Tartaglia, James
APPLICANT: No. 5766599ton, Elizabeth K.
APPLICANT: Riviere, Michel
APPLICANT: de Taisne, Charles
APPLICANT: Limbach, Keith J.
APPLICANT: Johnson, Gerard P.
APPLICANT: Pincus, Steven E.
APPLICANT: Cox, William I.
APPLICANT: Audonnet, Jean-Christophe Francis
TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE
TITLE OF INVENTION: STRAIN
NUMBER OF SEQUENCES: 467
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford
ADDRESSER: c/o William S. Frommer
STREET: 530 Fifth Avenue

CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,101
FILING DATE: 01-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2740
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 216:
SEQUENCE CHARACTERISTICS:
LENGTH: 389 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-458-101-216

Query Match 100.0%; Score 61; DB 1; Length 389;
Best Local Similarity 100.0%; Pred. No. 0.00086;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PLSIFSRIGDP 12
Db 149 PLSIFSRIGDP 160

RESULT 12
US-08-458-101-219
Sequence 219, Application US/08458101
Patent No. 5766599
GENERAL INFORMATION:
APPLICANT: Paoletti, Enzo
APPLICANT: Perkins, Marion E.
APPLICANT: Taylor, Jill
APPLICANT: Tartaglia, James
APPLICANT: No. 5766599ton, Elizabeth K.
APPLICANT: Riviere, Michel
APPLICANT: de Taisne, Charles
APPLICANT: Limbach, Keith J.
APPLICANT: Johnson, Gerard P.
APPLICANT: Pincus, Steven E.
APPLICANT: Cox, William I.
APPLICANT: Audonnet, Jean-Christophe Francis
TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE
TITLE OF INVENTION: STRAIN
NUMBER OF SEQUENCES: 467
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford
ADDRESSER: c/o William S. Frommer
STREET: 530 Fifth Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/458,101

US-08-486-099-106
FILING DATE: 01-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2740
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 219:
SEQUENCE CHARACTERISTICS:
LENGTH: 389 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-458-101-219
Query Match 100.0%; Score 61; DB 1; Length 389;
Best Local Similarity 100.0%; Pred. No. 0.00086;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 1 PLSSIFSRIGDP 12
DB 149 PLSSIFSRIGDP 160
RESULT 13
US-08-486-099-106
Sequence 106, Application US/08486099
Patent No. 6013263
GENERAL INFORMATION:
APPLICANT: Bolognesi, Dani P.
APPLICANT: Matthews, Thomas J.
APPLICANT: Wild, Carl T.
APPLICANT: Barney, Shawn O.
APPLICANT: Lambert, Dennis M.
APPLICANT: Petleway, Stephen R.
APPLICANT: Langlois, Alphonse J.
TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HEPATITIS
TITLE OF INVENTION: B VIRUS TRANSMISSION
NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,099
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-031
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 106:
SEQUENCE CHARACTERISTICS:
LENGTH: 389 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein

US-08-486-099-106
Query Match 100.0%; Score 61; DB 3; Length 389;
Best Local Similarity 100.0%; Pred. No. 0.00086;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 1 PLSSIFSRIGDP 12
DB 149 PLSSIFSRIGDP 160
RESULT 14
US-08-360-107A-116
Sequence 116, Application US/08360107A
Patent No. 6017336
GENERAL INFORMATION:
APPLICANT: Bolognesi, Dani P.
APPLICANT: Matthews, Thomas J.
APPLICANT: Wild, Carl T.
APPLICANT: Barney, Shawn O.
APPLICANT: Lambert, Dennis M.
APPLICANT: Petleway, Stephen R.
APPLICANT: Langlois, Alphonse J.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
TITLE OF INVENTION: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
TITLE OF INVENTION: TRANSMISSION
NUMBER OF SEQUENCES: 149
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/360,107A
FILING DATE: 20-DEC-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-013
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 116:
SEQUENCE CHARACTERISTICS:
LENGTH: 389 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-360-107A-116
Query Match 100.0%; Score 61; DB 3; Length 389;
Best Local Similarity 100.0%; Pred. No. 0.00086;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 1 PLSSIFSRIGDP 12
DB 149 PLSSIFSRIGDP 160
RESULT 15
US-08-484-223B-106
Sequence 106, Application US/08484223B
Patent No. 6020459

```

; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Letteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; NUMBER OF SEQUENCES: 245
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,223B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELETYPE: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 106:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 389 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-484-223B-106

```

```

Query Match          100.0%; Score 61; DB 3; Length 389;
Best Local Similarity 100.0%; Pred. No. 0.00066;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 PLSIFSRIGDP 12
        |||||
Db      149 PLSIFSRIGDP 160

```

Search completed: February 3, 2003, 09:27:10
 Job time : 0.731707 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using BW model

Run on: February 3, 2003, 09:25:53 ; Search time 0.439024 Seconds
(without alignments)
551.547 Million cell updates/sec

Title: US-09-890-752a-20

Sequence: 1 PLSISFRIGDP 12

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 122226 seqs, 20178551 residues

Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PubliSearch Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PTI_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCYUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61	100.0	281	9	US-09-247-890-12
2	59	96.7	174	10	US-09-818-066-55
3	59	96.7	389	9	US-09-821-877-2
4	54	88.5	170	10	US-09-818-066-49
5	54	88.5	174	10	US-09-818-066-46
6	54	88.5	174	10	US-09-818-066-54
7	54	88.5	174	10	US-09-818-066-56
8	54	88.5	281	9	US-09-247-890-10
9	54	88.5	400	9	US-10-209-284-3
10	52	85.2	174	10	US-09-818-066-50
11	52	85.2	174	10	US-09-818-066-58
12	49	80.3	174	10	US-09-818-066-53
13	46	75.4	174	10	US-09-818-066-45
14	46	75.4	174	10	US-09-818-066-47
15	44	72.1	174	10	US-09-818-066-37
16	44	72.1	174	10	US-09-818-066-37
17	43	70.5	174	10	US-09-818-066-57
18	40	65.6	64	10	US-09-917-340-85
19	40	65.6	174	10	US-09-818-066-42

20	40	65.6	174	10	US-09-818-066-62	Sequence 62, Appl
21	38	62.3	174	10	US-09-818-066-38	Sequence 38, Appl
22	38	62.3	174	10	US-09-818-066-40	Sequence 40, Appl
23	37	62.3	174	10	US-09-818-066-41	Sequence 41, Appl
24	37	60.7	174	10	US-09-818-066-59	Sequence 59, Appl
25	36	59.0	55	10	US-09-879-257A-45	Sequence 45, Appl
26	36	59.0	174	10	US-09-818-066-35	Sequence 35, Appl
27	36	59.0	174	10	US-09-818-066-36	Sequence 36, Appl
28	36	59.0	174	10	US-09-818-066-36	Sequence 36, Appl
29	35	57.4	96	10	US-09-864-761-39796	Sequence 39796, A
30	35	57.4	173	9	US-10-051-643-207	Sequence 207, Appl
31	35	57.4	174	10	US-09-818-066-39	Sequence 39, Appl
32	35	57.4	690	10	US-09-815-242-12460	Sequence 12460, A
33	35	57.4	703	10	US-09-815-242-5610	Sequence 5610, A
34	35	57.4	705	9	US-09-738-626-4055	Sequence 4055, Ap
35	35	57.4	743	10	US-09-771-161A-164	Sequence 164, App
36	35	57.4	743	10	US-09-771-161A-254	Sequence 254, App
37	35	57.4	1036	10	US-09-771-161A-255	Sequence 255, App
38	35	57.4	1036	10	US-09-771-161A-256	Sequence 256, App
39	34	55.7	64	10	US-09-917-340-84	Sequence 84, Appl
40	34	55.7	119	10	US-09-825-299-799	Sequence 799, App
41	34	55.7	174	10	US-09-818-066-43	Sequence 43, Appl
42	34	55.7	174	10	US-09-818-066-44	Sequence 44, Appl
43	34	55.7	284	9	US-09-895-913A-102	Sequence 102, App
44	34	55.7	312	10	US-09-801-368-260	Sequence 260, Appl
45	34	55.7	478	9	US-10-156-634A-2	Sequence 2, Appl1

ALIGNMENTS

```
RESULT 1
US-09-247-890-12
; Sequence 12, Application US/09247890
; Publication NO. US20020198162A1
GENERAL INFORMATION:
; APPLICANT: Punnonen, Juhana
; APPLICANT: Baes, Steven H.
; APPLICANT: Whalen, Robert Gerald
; APPLICANT: Howard, Russell P.C.
; APPLICANT: Stemmer, Willem P.C.
; APPLICANT: Maxygen, Inc.
; TITLE OF INVENTION: Antigen Library Immunization
; FILE REFERENCE: 018097-028710US
; CURRENT APPLICATION NUMBER: US/09/247, 890
; EARLIER FILING DATE: 1998-02-10
; EARLIER APPLICATION NUMBER: US 60/074,294
; EARLIER FILING DATE: 1998-02-11
; EARLIER APPLICATION NUMBER: US 60/105,509
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 281
; TYPE: PRT
; ORGANISM: Hepatitis B virus
US-09-247-890-12
Query Match 100.0%; Score 61; DB 9; Length 281;
Best Local Similarity 100.0%; Pred. No. 0.00052;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy 1 PLSISFRIGDP 12
Db 41 PLSISFRIGDP 52
RESULT 2
US-09-818-066-55
; Sequence 55, Application US/09818066
; Patent No. US20020032307A1
; GENERAL INFORMATION:
; APPLICANT: Shuping Tong et al.
```

TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR
NUMBER OF SEQUENCES: 75
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/818,066
FILING DATE: 27-Mar-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/683,262
FILING DATE: 18-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 31,819
REFERENCE/DOCKET NUMBER: 00786/287002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 174 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 55:
US-09-818-066-55

Query Match 96.7%; Score 59; DB 10; Length 174;
Best Local Similarity 91.7%; Pred. No. 0.0073;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSIFSRIGDP 12
DB 160 PLSIFSRIGDP 171

RESULT 3
US-09-821-877-2
Sequence 2, Application US/09821877
Patent No. US20020177124A1
GENERAL INFORMATION:
APPLICANT: Abbott Laboratories
APPLICANT: Coleman, Paul F.
APPLICANT: Mushawar, Isa K.
TITLE OF INVENTION: Hepatitis B Virus Surface Antigen Mutant
TITLE OF INVENTION: And Methods Of Detection Thereof
FILE REFERENCE: 6794, US-01
CURRENT APPLICATION NUMBER: US/09/821,877
CURRENT FILING DATE: 2000-03-03
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PaeSeq for Windows Version 4.0
SEQ ID NO: 2
LENGTH: 389
TYPE: PRT
ORGANISM: Hepatitis B Virus
US-09-821-877-2

Query Match 96.7%; Score 59; DB 9; Length 389;
Best Local Similarity 91.7%; Pred. No. 0.0017;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSIFSRIGDP 12

DB 149 PLSIFSRIGDP 160

RESULT 4
US-09-818-066-49
Sequence 49, Application US/09818066
Patent No. US20020032307A1
GENERAL INFORMATION:
APPLICANT: Shuping Tong et al.
TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR
NUMBER OF SEQUENCES: 75
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/818,066
FILING DATE: 27-Mar-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/683,262
FILING DATE: 18-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 31,819
REFERENCE/DOCKET NUMBER: 00786/287002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 170 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 49:
US-09-818-066-49

Query Match 88.5%; Score 54; DB 10; Length 170;
Best Local Similarity 83.3%; Pred. No. 0.0058;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLSIFSRIGDP 12
DB 156 PLSIFSRIGDP 167

RESULT 5
US-09-818-066-46
Sequence 46, Application US/09818066
Patent No. US20020032307A1
GENERAL INFORMATION:
APPLICANT: Shuping Tong et al.
TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR
NUMBER OF SEQUENCES: 75
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/818,066
 FILING DATE: 27-Mar-2001
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/683,262
 FILING DATE: 18-JUL-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Fraser, Janis K.
 REGISTRATION NUMBER: 31,819
 REFERENCE/DOCKET NUMBER: 00786/287002
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 542-5070
 TELEFAX: (617) 542-8906
 TELEX: 200154
 INFORMATION FOR SEQ ID NO: 46:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 174 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 46:
 US-09-818-066-46

Query Match 88.5%; Score 54; DB 10; Length 174;
 Best Local Similarity 83.3%; Pred. No. 0.006;
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 PLSISFRTGDP 12
 Db 160 PLSISFRTGDP 171

RESULT 6
 US-09-818-066-48
 Sequence 48, Application US/09818066
 Patent No. US20020032307A1
 GENERAL INFORMATION:
 APPLICANT: Shuping Tong et al.
 TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR
 NUMBER OF SEQUENCES: 75
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson P.C.
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: U.S.A.
 ZIP: 02110-2804
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/818,066
 FILING DATE: 27-Mar-2001
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/683,262
 FILING DATE: 18-JUL-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Fraser, Janis K.
 REGISTRATION NUMBER: 31,819
 REFERENCE/DOCKET NUMBER: 00786/287002
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 542-5070
 TELEFAX: (617) 542-8906
 TELEX: 200154
 INFORMATION FOR SEQ ID NO: 48:

SEQUENCE CHARACTERISTICS:
 LENGTH: 174 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 48:
 US-09-818-066-48

Query Match 88.5%; Score 54; DB 10; Length 174;
 Best Local Similarity 83.3%; Pred. No. 0.006;
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 PLSISFRTGDP 12
 Db 160 PLSISFRTGDP 171

RESULT 7
 US-09-818-066-56
 Sequence 56, Application US/09818066
 Patent No. US20020032307A1
 GENERAL INFORMATION:
 APPLICANT: Shuping Tong et al.
 TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR
 NUMBER OF SEQUENCES: 75
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson P.C.
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: U.S.A.
 ZIP: 02110-2804
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/818,066
 FILING DATE: 27-Mar-2001
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/683,262
 FILING DATE: 18-JUL-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Fraser, Janis K.
 REGISTRATION NUMBER: 31,819
 REFERENCE/DOCKET NUMBER: 00786/287002
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 542-5070
 TELEFAX: (617) 542-8906
 TELEX: 200154
 INFORMATION FOR SEQ ID NO: 56:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 174 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 56:
 US-09-818-066-56

Query Match 88.5%; Score 54; DB 10; Length 174;
 Best Local Similarity 100.0%; Pred. No. 0.006;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 PLSISFRTGDP 11
 Db 160 PLSISFRTGDP 170

RESULT 8
 US-09-247-890-10
 Sequence 10, Application US/09247890

```
; Publication No. US20020198162A1
; GENERAL INFORMATION:
; APPLICANT: Punnonen, Juha
; APPLICANT: Babs, Steven H.
; APPLICANT: Whalen, Robert Gerald
; APPLICANT: Howard, Russell P.C.
; APPLICANT: Stemmer, William P.C.
; TITLE OF INVENTION: Antigen Library Immunization
; FILE REFERENCE: 018097-028710US
; CURRENT APPLICATION NUMBER: US/09/247,690
; CURRENT FILING DATE: 1998-02-10
; EARLIER APPLICATION NUMBER: US 60/074,294
; EARLIER FILING DATE: 1998-02-11
; EARLIER APPLICATION NUMBER: US 60/105,509
; EARLIER FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 281
; TYPE: PRT
; ORGANISM: Hepatitis B virus
; US-09-247-890-10

Query Match      88.5%; Score 54; DB 9; Length 281;
Best Local Similarity 83.3%; Pred. No. 0.01;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLSIFSRIGDP 12
Db 41 PLSIFSRIGDP 52

RESULT 9
US-10-209-264-3
; Sequence 3, Application US/10209264
; Publication No. US20030003111A1
; GENERAL INFORMATION:
; APPLICANT: Oon, Chong Jin
;           Lim, Gek Keow
;           Chen, Wei Ning
;           Zhao, Yi
; TITLE OF INVENTION: A MUTANT HUMAN HEPATITIS B VIRAL STRAIN AND
;           USES THEREOF
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ladass & Parry
; STREET: 26 West 61 Street
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10023
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/209,264
; FILING DATE: 31-Jul-2002
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/SG98/00046
; FILING DATE: 19-JAN-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Mass, Clifford J.
; REGISTRATION NUMBER: 30,086
; REFERENCE/DOCKET NUMBER: U-013109-7
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 708-1800
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 400 amino acids
```

```
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
; US-10-209-264-3

Query Match      88.5%; Score 54; DB 9; Length 400;
Best Local Similarity 83.3%; Pred. No. 0.015;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLSIFSRIGDP 12
Db 160 PLSIFSRIGDP 171

RESULT 10
US-09-818-066-50
; Sequence 50, Application US/09818066
; Patent No. US20020032307A1
; GENERAL INFORMATION:
; APPLICANT: Shuping Tong et al.
; TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; MEDIUM TYPE: Floppy disk
; COMPUTER READABLE FORM:
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/818,066
; FILING DATE: 27-Mar-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/683,262
; FILING DATE: 18-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Frazer, Janis K.
; REGISTRATION NUMBER: 31,819
; REFERENCE/DOCKET NUMBER: 00786/287002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 174 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 50:
; US-09-818-066-50

Query Match      85.2%; Score 52; DB 10; Length 174;
Best Local Similarity 90.9%; Pred. No. 0.014;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PLSIFSRIGDP 12
Db 161 PLSIFSRIGDP 171

RESULT 11
US-09-818-066-58
; Sequence 58, Application US/09818066
; Patent No. US20020032307A1
; GENERAL INFORMATION:
```


APPLICANT: Shuping Tong et al.
TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR
NUMBER OF SEQUENCES: 75
CORRESPONDENCE ADDRESSES:
ADDRESSER: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/818,066
FILING DATE: 27-Mar-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/683,262
FILING DATE: 18-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 31,819
REFERENCE/DOCKET NUMBER: 00786/287002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 174 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 58:
US-09-818-066-58
Query Match 85.2%; Score 52; DB 10; Length 174;
Best Local Similarity 90.9%; Pred. No. 0.014;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 LSSIFSRIGDP 12
Db 161 ISSIFSRIGDP 171
RESULT 12
US-09-818-066-53
Sequence 53, Application US/09818066
Patent No. US20020032307A1
GENERAL INFORMATION:
APPLICANT: Shuping Tong et al.
TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR
NUMBER OF SEQUENCES: 75
CORRESPONDENCE ADDRESSES:
ADDRESSER: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/818,066
FILING DATE: 27-Mar-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/683,262
FILING DATE: 18-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 31,819
REFERENCE/DOCKET NUMBER: 00786/287002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 174 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 53:
US-09-818-066-53
Query Match 80.3%; Score 49; DB 10; Length 174;
Best Local Similarity 81.8%; Pred. No. 0.049;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 2 LSSIFSRIGDP 12
Db 161 ISSIFSRIGDP 171
RESULT 13
US-09-818-066-45
Sequence 45, Application US/09818066
Patent No. US20020032307A1
GENERAL INFORMATION:
APPLICANT: Shuping Tong et al.
TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR
NUMBER OF SEQUENCES: 75
CORRESPONDENCE ADDRESSES:
ADDRESSER: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/818,066
FILING DATE: 27-Mar-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/683,262
FILING DATE: 18-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 31,819
REFERENCE/DOCKET NUMBER: 00786/287002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 174 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-09-818-066-45
Query Match 75.4%; Score 46; DB 10; Length 174;

Best Local Similarity 75.0%; Pred. No. 0.18;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 LSSISFRIGDP 12
DB 160 PISISSRTGDP 171

RESULT 14

US-09-818-066-47
Sequence 47, Application US/09818066
Patent No. US20020032307A1

GENERAL INFORMATION:

APPLICANT: Shuping Tong et al.
TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR
NUMBER OF SEQUENCES: 75
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/818,066
FILING DATE: 27-Mar-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/683,262
FILING DATE: 18-JUL-1996

ATTORNEY/AGENT INFORMATION:

NAME: Fraser, Janis K.
REGISTRATION NUMBER: 31,819
REFERENCE/DOCKET NUMBER: 00786/287002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154

INFORMATION FOR SEQ ID NO: 47:

SEQUENCE CHARACTERISTICS:
LENGTH: 174 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 47:

US-09-818-066-47

Query Match 75.4%; Score 46; DB 10; Length 174;

Best Local Similarity 75.0%; Pred. No. 0.18;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 LSSISFRIGDP 12
DB 160 PISISSRTGDP 171

RESULT 15

US-09-818-066-37
Sequence 37, Application US/09818066
Patent No. US20020032307A1

GENERAL INFORMATION:

APPLICANT: Shuping Tong et al.
TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR
NUMBER OF SEQUENCES: 75
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston

STATE: Massachusetts

COUNTRY: U.S.A.
ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/818,066
FILING DATE: 27-Mar-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/683,262
FILING DATE: 18-JUL-1996

ATTORNEY/AGENT INFORMATION:

NAME: Fraser, Janis K.
REGISTRATION NUMBER: 31,819
REFERENCE/DOCKET NUMBER: 00786/287002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154

INFORMATION FOR SEQ ID NO: 37:

SEQUENCE CHARACTERISTICS:
LENGTH: 174 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 37:

US-09-818-066-37

Query Match 72.1%; Score 44; DB 10; Length 174;

Best Local Similarity 81.8%; Pred. No. 0.41;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 LSSISFRIGDP 12
DB 161 ISSISSRTGDP 171

Search completed: February 3, 2003, 09:35:29
Job time: 0.439024 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: February 3, 2003, 09:24:27 ; Search time 0.982578 Seconds
(without alignments)
988.252 Million cell updates/sec

Title: US-09-890-752A-20
Perfect score: 61
Sequence: 1 PLSIFSRIDGP 12

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 404799 seqs, 80919614 residues
Total number of hits satisfying chosen parameters: 404799

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents AA New:
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	61	100.0	12	US-10-211-088-302	Sequence 302, App
2	59	96.7	64	US-10-326-908-19	Sequence 19, App
3	59	96.7	67	US-10-326-908-21	Sequence 21, App
4	54	88.5	400	PCT-US02-33263-46	Sequence 46, App
5	54	88.5	400	US-10-287-994-46	Sequence 46, App
6	48	78.7	10	PCT-US02-13609-23	Sequence 23, App
7	39	63.9	399	PCT-US02-40225-3144	Sequence 3144, App
8	39	63.9	399	US-10-320-797-3144	Sequence 3144, App
9	37	60.7	533	US-10-092-411A-4053	Sequence 4053, App
10	35	57.4	53	PCT-US02-32727-2247	Sequence 2247, App
11	35	57.4	53	US-10-057-498-2247	Sequence 2247, App
12	35	57.4	70	PCT-US02-32727-19192	Sequence 19192, App
13	35	57.4	70	US-10-057-498-19192	Sequence 19192, App
14	35	57.4	144	PCT-US02-32727-23079	Sequence 23079, App
15	35	57.4	144	US-10-057-498-23079	Sequence 23079, App
16	35	57.4	199	US-09-724-676-70237	Sequence 70237, App
17	35	57.4	244	US-09-724-676-70238	Sequence 70238, App
18	35	57.4	244	US-09-724-676-70237	Sequence 70237, App
19	35	57.4	244	US-09-724-676-70238	Sequence 70238, App
20	35	57.4	244	US-09-724-676-70237	Sequence 70237, App
21	35	57.4	309	PCT-US02-34321-5	Sequence 5, App
22	35	57.4	309	PCT-US02-34321-3	Sequence 3, App
23	35	57.4	322	PCT-US02-32727-29962	Sequence 29962, App
24	35	57.4	687	US-09-724-676-70230	Sequence 70230, App
25	35	57.4	687	US-09-724-676-70231	Sequence 70231, App
26	35	57.4	687	US-09-724-676A-70230	Sequence 70230, App

27	35	57.4	687	5	US-09-724-676A-70231	Sequence 70231, A
28	35	57.4	696	5	US-09-724-676-70232	Sequence 70232, A
29	35	57.4	696	5	US-09-724-676-70233	Sequence 70233, A
30	35	57.4	696	5	US-09-724-676A-70232	Sequence 70232, A
31	35	57.4	696	5	US-09-724-676A-70233	Sequence 70233, A
32	35	57.4	741	1	PCT-US02-32727-1231	Sequence 1231, App
33	35	57.4	741	6	US-10-057-498-1231	Sequence 1231, App
34	35	57.4	1037	5	US-09-724-676-70223	Sequence 70223, A
35	35	57.4	1037	5	US-09-724-676-70223	Sequence 70223, A
36	35	57.4	1047	5	US-09-724-676-70234	Sequence 70234, A
37	35	57.4	1047	5	US-09-724-676-70239	Sequence 70239, A
38	35	57.4	1047	5	US-09-724-676-70234	Sequence 70234, A
39	35	57.4	1047	5	US-09-724-676A-70239	Sequence 70239, A
40	35	57.4	1069	5	US-09-724-676-70240	Sequence 70240, A
41	35	57.4	1069	5	US-09-724-676-70241	Sequence 70241, A
42	35	57.4	1069	5	US-09-724-676A-70240	Sequence 70240, A
43	35	57.4	1069	5	US-09-724-676A-70241	Sequence 70241, A
44	35	57.4	1075	5	US-09-724-676-70224	Sequence 70224, A
45	35	57.4	1075	5	US-09-724-676-70225	Sequence 70225, A

ALIGNMENTS

RESULT 1
US-10-211-088-302
; Sequence 302, Application US/10211088
; GENERAL INFORMATION:
; APPLICANT: Bright, Gary R.
; APPLICANT: Premkumar, D. David
; APPLICANT: Chen, Yih-Tai
; TITLE OF INVENTION: Novel Fusion Proteins And Assays For Molecular Binding
; FILE REFERENCE: 01-1022-US
; CURRENT APPLICATION NUMBER: US/10/211,088
; PRIOR FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/309,395
; PRIOR FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/341,589
; PRIOR FILING DATE: 2001-12-13
; NUMBER OF SEQ ID NOS: 366
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 302
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Protein-derived transport peptide
US-10-211-088-302

Query Match 100.0%; Score 61; DB 6; Length 12;
Best local Similarity 100.0%; Pred. No. 4.6e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query 1 PLSIFSRIDGP 12
DB 1 PLSIFSRIDGP 12
US-10-326-908-19
; Sequence 19, Application US/10326908
; GENERAL INFORMATION:
; APPLICANT: Christopher JONES
; APPLICANT: Andrew BACON
; APPLICANT: Gill DOUCE
; APPLICANT: Mark PAGE
; TITLE OF INVENTION: Designing Immunogens
; FILE REFERENCE: 117-433 / N78027D
; CURRENT APPLICATION NUMBER: US/10/326,908
; CURRENT FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: US 10/110,036
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: PCT/GB00/03857

```

; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: GB 0007789.1
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: GB 9923902.2
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: MS Word
; SEQ ID NO 19
; LENGTH: 64
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: synthetic peptide immunogen
US-10-326-908-19

Query Match
Best Local Similarity 96.7%; Score 59; DB 6; Length 64;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSIFSRIGDP 12
Db 48 PISIFSRIGDP 59

RESULT 3
US-10-326-908-21
; Sequence 21, Application US/10326908
; GENERAL INFORMATION:
; APPLICANT: Christopher JONES
; APPLICANT: Andrew BACON
; APPLICANT: Gill DOUCE
; APPLICANT: Mark PAGE
; TITLE OF INVENTION: Designing Immunogens
; FILE REFERENCE: 117-433 / N78027D
; CURRENT FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: US 10/110,036
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: PCT/GB00/03857
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: GB 0007789.1
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: GB 9923902.2
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: MS Word
; SEQ ID NO 21
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: synthetic peptide immunogen
US-10-326-908-21

Query Match
Best Local Similarity 96.7%; Score 59; DB 6; Length 67;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSIFSRIGDP 12
Db 48 PISIFSRIGDP 59

RESULT 4
PCT-US02-32263-46
; Sequence 46, Application PC/TUS0232263
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: Deftrees, Shawn
; APPLICANT: Zopf, David
; APPLICANT: Bayer, Robert
; APPLICANT: Bower, Caryn
; APPLICANT: Hakes, David
```

```

; APPLICANT: Chen, Xi
; TITLE OF INVENTION: REMODELING AND GLYCOCONJUGATION OF PEPTIDES
; FILE REFERENCE: 040853-01-5050WO
; CURRENT APPLICATION NUMBER: PCT/US02/32263
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US 60/328,523
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/344,692
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/334,233
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/334,301
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/387,292
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 60/391,777
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46
; LENGTH: 400
; TYPE: PRT
; ORGANISM: Hepatitis B virus
PCT-US02-32263-46
```

```

Query Match
Best Local Similarity 88.5%; Score 54; DB 1; Length 400;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PLSIFSRIGDP 12
Db 160 PISIFSRIGDP 171
```

```

RESULT 5
US-10-287-994-46
; Sequence 46, Application US/10287994
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: Deftrees, Shawn
; APPLICANT: Zopf, David
; APPLICANT: Bayer, Robert
; APPLICANT: Bower, Caryn
; APPLICANT: Hakes, David
; APPLICANT: Chen, Xi
; TITLE OF INVENTION: REMODELING AND GLYCOCONJUGATION OF PEPTIDES
; FILE REFERENCE: 040853-01-5052-00
; CURRENT APPLICATION NUMBER: US/10/287,994
; CURRENT FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: US 60/328,523
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/344,692
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/387,292
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 60/391,777
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46
; LENGTH: 400
```

TYPE: PRT
ORGANISM: Hepatitis B virus
US-10-287-994-46

Query Match 88.5%; Score 54; DB 6; Length 400;
Best Local Similarity 83.3%; Pred. No. 0.036;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PLSIFSRIGD 12
Db 160 PLSIFSRIGD 171

RESULT 6
PCT-US02-13609-23
Sequence 23, Application PC/TUS0213609

GENERAL INFORMATION:
APPLICANT: Targeted Genetics Corporation
APPLICANT: Emerald Gene Systems, LTD
APPLICANT: Harvie, Pierre
APPLICANT: Paul, Ralph
APPLICANT: O'Mahony, Daniel J.
TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES
FILE REFERENCE: 226272005340
CURRENT APPLICATION NUMBER: PCT/US02/13609
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/287,786
NUMBER OF SEQ ID NOS: 45
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 23
LENGTH: 10
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic Construct
PCT-US02-13609-23

Query Match 78.7%; Score 48; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSIFSRIG 10
Db 1 PLSIFSRIG 10

RESULT 7
PCT-US02-40225-3144
Sequence 3144, Application PC/TUS0240225
GENERAL INFORMATION:
APPLICANT: Elixir Pharmaceuticals, Inc.
APPLICANT: Eroshkin, Alexey M.
APPLICANT: Zamudio, Carlos
TITLE OF INVENTION: IDENTIFICATION OF ESSENTIAL GENES OF CRYPTOCOCCUS NEOFORMANS AND
FILE REFERENCE: 10182-021-228
CURRENT APPLICATION NUMBER: PCT/US02/40225
CURRENT FILING DATE: 2002-12-17
PRIOR APPLICATION NUMBER: 60/341,261
PRIOR FILING DATE: 2001-12-17
NUMBER OF SEQ ID NOS: 3361
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3144
LENGTH: 399
TYPE: PRT
ORGANISM: Cryptococcus neoformans
PCT-US02-40225-3144

Query Match 63.9%; Score 39; DB 1; Length 399;
Best Local Similarity 70.0%; Pred. No. 25;

Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
Qy 2 LSAVSRIGD 11
Db 213 LSAVSRIGD 222

RESULT 8
US-10-320-797-3144
Sequence 3144, Application US/10320797

GENERAL INFORMATION:
APPLICANT: Eroshkin, Alexey M.
APPLICANT: Zamudio, Carlos
TITLE OF INVENTION: IDENTIFICATION OF ESSENTIAL GENES OF CRYPTOCOCCUS NEOFORMANS AND
FILE REFERENCE: 10182-021-999
CURRENT APPLICATION NUMBER: US/10/320,797
CURRENT FILING DATE: 2002-12-16
PRIOR APPLICATION NUMBER: 60/341,261
PRIOR FILING DATE: 2001-12-17
NUMBER OF SEQ ID NOS: 3361
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3144
LENGTH: 399
TYPE: PRT
ORGANISM: Cryptococcus neoformans
US-10-320-797-3144

Query Match 63.9%; Score 39; DB 6; Length 399;
Best Local Similarity 70.0%; Pred. No. 25;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSAVSRIGD 11
Db 213 LSAVSRIGD 222

RESULT 9
US-10-092-411A-4053
Sequence 4053, Application US/10092411A
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: 032796-101
CURRENT APPLICATION NUMBER: US/10/092,411A
CURRENT FILING DATE: 2002-03-07
PRIOR APPLICATION NUMBER: US 09/134,001
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5676
SEQ ID NO 4053
LENGTH: 533
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-10-092-411A-4053

Query Match 60.7%; Score 37; DB 6; Length 533;
Best Local Similarity 60.0%; Pred. No. 82;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSIFSRIG 10
Db 98 PLSIFSRIG 107

RESULT 10
PCT-US02-32727-2247
Sequence 2247, Application PC/TUS0232727
GENERAL INFORMATION:

/ CURRENT APPLICATION NUMBER: PCT/US02/32727
 / CURRENT FILING DATE: 2002-10-11
 / NUMBER OF SEQ ID NOS: 30992
 / SEQ ID NO 23079
 / LENGTH: 144
 / TYPE: PRT
 / ORGANISM: Propioni acnes
 PCT-US02-32727-23079

Query Match 57.4%; Score 35; DB 1; Length 144;
 Best Local Similarity 45.5%; Pred. No. 51;
 Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 LSSIFSRIGDP 12
 DB 24 LAAVFKSLGDP 34

RESULT 15
 US-10-057-498-23079
 / Sequence 23079, Application US/10057498
 / GENERAL INFORMATION:
 / APPLICANT: Mitcham, Jennifer
 / APPLICANT: Skeiky, Yashir
 / APPLICANT: Persing, David
 / TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes
 / FILE REFERENCE: 210121.514
 / CURRENT APPLICATION NUMBER: US/10/057,498
 / CURRENT FILING DATE: 2001-04-20
 / NUMBER OF SEQ ID NOS: 29212
 / SEQ ID NO 23079
 / LENGTH: 144
 / TYPE: PRT
 / ORGANISM: Propioni acnes
 US-10-057-498-23079

Query Match 57.4%; Score 35; DB 6; Length 144;
 Best Local Similarity 45.5%; Pred. No. 51;
 Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 LSSIFSRIGDP 12
 DB 24 LAAVFKSLGDP 34

Search completed: February 3, 2003, 09:35:01
 Job time : 1.98258 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 09:21:57 ; Search time 0.794425 Seconds
(without alignments)
1452.136 Million cell updates/sec

Title: US-09-890-752a-20

Sequence: 1 PLSIFSRIGDP 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 73:*
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61	100.0	378	2 S41869	surface antigen -
2	61	100.0	389	1 SAVLAH	large surface anti-
3	61	100.0	389	2 S20749	surface antigen -
4	59	96.7	389	1 SAVLBH	large surface anti-
5	59	96.7	389	1 SAVLAI	large surface anti-
6	59	96.7	389	1 SAVLAI	large surface anti-
7	59	96.7	389	2 S47407	surface antigen -
8	59	96.7	389	2 S32202	large surface anti-
9	54	88.5	384	2 T13474	large surface anti-
10	54	88.5	384	2 T13469	large surface anti-
11	54	88.5	389	2 S20745	large surface anti-
12	54	88.5	400	1 SAVLA	large surface anti-
13	54	88.5	400	2 S35528	surface antigen -
14	54	88.5	445	2 S43482	surface antigen -
15	47	77.0	389	2 S67506	large surface anti-
16	47	77.0	389	2 S41870	surface antigen -
17	45	73.8	389	2 S20753	surface antigen -
18	43	70.5	358	2 B81194	conserved hypochet
19	43	70.5	358	2 A81831	ubiquitinone biosynt
20	41	67.2	502	2 C87709	hypothetical prote
21	40	65.6	389	1 SAVLCP	surface antigen -
22	40	65.6	389	2 S41871	hypothetical prote
23	39	63.9	399	2 AF2349	hypothetical large
24	39	63.9	557	2 S42226	hypothetical large
25	38	62.3	278	2 AF2895	hypothetical prote
26	38	62.3	315	2 A86807	hypothetical prote
27	38	62.3	389	1 SAVLAI	large surface anti-
28	38	62.3	389	1 SAVLAI	large surface anti-
29	38	62.3	449	2 F86852	amino acid pernease

ALIGNMENTS

RESULT 1

S41869 surface antigen - hepatitis B virus (subtype ayw, patient A) (fragment)

N/Alternate names: envelope protein; HBs antigen

N/Contains: surface antigen pre-S1 (large envelope protein); surface antigen pre-S2 (mid-

C/Species: hepatitis B virus, HBV

A/Variety: subtype ayw, patient A

C/Date: 06-Jan-1995 #sequence_revision 01-Nov-1996 #text_change 26-Aug-1999

R/Lat, M.E.; Mazzoleni, A.P.; Balestrieri, A.

submitted to the EMBL Data Library, January 1994

A/Description: Sequence analysis of HBV genomes isolated from patients with HBsAg chronic

A/Reference number: S41869

A/Accession: S41869

A/Molecule type: DNA

A/Residues: 1-378 <LAI>

A/Cross-references: EMBL:X77309; NID:G452611; PIDN:CA54515.1; PID:G452612

A/Experimental source: subtype ayw, patient A

A/Genetics:

C/Superfamily: hepatitis B virus surface antigen

C/Keywords: surface antigen

F/1-378/Product: surface antigen pre-S1 (large envelope protein) #status predicted <PS1>

F/1-108/Domain: pre-S1 domain #status predicted <PR1>

F/109-378/Product: surface antigen pre-S2 (middle envelope protein) #status predicted <PS2>

F/109-163/Domain: pre-S2 domain #status predicted <PR2>

F/164-378/Product: surface antigen S (small envelope protein) #status predicted <PSD>

Query Match

Best local similarity 100.0%; Score 61; DB 2; Length 378;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSIFSRIGDP 12

DB 149 PLSIFSRIGDP 160

RESULT 2

SAVLAH

large surface antigen - hepatitis B virus (subtype ayw)

N/Contains: major surface antigen; middle surface antigen

C/Species: hepatitis B virus, HBV

C/Date: 18-Dec-1981 #sequence_revision 18-Dec-1991 #text_change 07-May-1999

C/Accession: A03703; J02064; F00591

R/Galbert, F.; Mandart, B.; Pitoussi, F.; Riollais, P.; Charney, P.

Nature 281, 646-650, 1979

A/Title: Nucleotide sequence of the hepatitis B virus genome (subtype ayw) in B. coli.

A/Reference number: A93214; MUID:81012091; PMID:199227

A/Accession: A03703

A/Molecule type: DNA

A/Residues: 1-389 <GAL>

A/Cross-references: GB:J02203

R/Norder, H.; Hammas, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnus, J. Gen. Virol. 74, 1341-1348, 1993
 A/Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
 A/Accession: J02044; MUID:93329382; PMID:8336122
 A/Molecule type: DNA
 A/Residues: 164-389 <NOR>
 A/Experimental source: genogroup D, subtype ayw3
 R/Norder, H.; Courouce, A.M.; Magnus, L.O.
 J. Gen. Virol. 73, 3141-3145, 1992
 A/Title: Molecular basis of hepatitis B virus serotype variations within the four major
 A/Reference number: P00453; MUID:93107848; PMID:1469353
 A/Accession: P00591
 A/Molecule type: DNA
 A/Residues: 264-343 <NO2>
 A/Experimental source: subtype ayw3, K11
 A/Genetic: S
 C/Superfamily: hepatitis B virus surface antigen
 C/Keywords: glycoprotein; surface antigen (gene pre-S2/S) #status predicted <DSA>
 F/109-389/Product: middle surface antigen (gene S) #status predicted <MSA>
 F/164-389/Product: major surface antigen (gene S) #status predicted
 F/4,112,166/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F/4,112,166/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 61; DB 1; Length 389;
 Best Local Similarity 100.0%; Pred. No. 0.0012;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSSIFSRIGDP 12
 DB 149 PLSSIFSRIGDP 160

RESULT 3
 S20749
 Surface antigen - hepatitis B virus (subtype ayw, patient CI)
 N/Alternative names: envelope protein, HBs antigen
 N/Contains: major surface antigen; middle surface antigen
 C/Species: hepatitis B virus, HBV
 A/Variety: subtype ayw, patient CI
 C/Date: 20-Feb-1995 #sequence_revision 01-Nov-1996 #text_change 26-Aug-1999
 C/Accession: S20749
 R/Alai, M.B.; Mazzolani, A.P.; Balestrieri, A.; Melis, A.; Porru, A.
 Submitted to the EMBL Data Library, March 1992
 A/Description: Sequence analysis of HBV genomes isolated from patients with HBsAg negative
 A/Reference number: S20745
 A/Accession: S20749
 A/Molecule type: DNA
 A/Residues: 1-389 <LAI>
 A/Cross-references: EMBL:X65258; NID:G59434; PIDN:CAA46553.1; PID:G59435
 A/Experimental source: subtype ayw, patient CI
 A/Note: typical methionine for the beginning of pre-S2 domain at position 109 is missing
 C/Genetic: S
 C/Superfamily: hepatitis B virus surface antigen
 C/Keywords: surface antigen
 F/1-389/Product: surface antigen pre-S1 (large envelope protein) #status predicted <PS1>
 F/164-389/Product: surface antigen S (small envelope protein) #status predicted <PS2>

Query Match 100.0%; Score 61; DB 2; Length 389;
 Best Local Similarity 100.0%; Pred. No. 0.0012;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSSIFSRIGDP 12
 DB 149 PLSSIFSRIGDP 160

RESULT 4
 SAV1BH
 Large surface antigen - hepatitis B virus (subtype ayw, strain PHB320)
 N/Contains: major surface antigen; middle surface antigen
 C/Species: hepatitis B virus, HBV

A/Note: host Homo sapiens (man)
 C/Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 28-Jul-2000
 C/Accession: A03704; P00585
 R/Bichko, V.; Pushko, P.; Dreilina, D.; Pumpen, P.; Gren, E.
 FEBS Lett. 185, 208-212, 1985
 A/Title: Subtype ayw variant of hepatitis B virus: DNA primary structure analysis.
 A/Reference number: A05237; MUID:85204397; PMID:396597
 A/Accession: A03704
 A/Molecule type: DNA
 A/Residues: 1-389 <BIC>
 A/Cross-references: GB:X02496; NID:962280; PIDN:CAB41701.1; PID:G4704321
 R/Norder, H.; Courouce, A.M.; Magnus, L.O.
 J. Gen. Virol. 73, 3141-3145, 1992
 A/Title: Molecular basis of hepatitis B virus serotype variations within the four major
 A/Reference number: P00453; MUID:93107848; PMID:1469353
 A/Accession: P00585
 A/Molecule type: DNA
 A/Residues: 264-343 <NOR>
 A/Experimental source: subtype ayw2, Tav
 C/Genetic: S
 C/Superfamily: hepatitis B virus surface antigen
 C/Keywords: glycoprotein; surface antigen (gene pre-S2/S) #status predicted <DSA>
 F/109-389/Product: middle surface antigen (gene S) #status predicted <MSA>
 F/164-389/Product: major surface antigen (gene S) #status predicted
 F/4,112,166/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 96.7%; Score 59; DB 1; Length 389;
 Best Local Similarity 91.7%; Pred. No. 0.0028;
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSSIFSRIGDP 12
 DB 149 PLSSIFSRIGDP 160

RESULT 5
 SAV1AJ
 Large surface antigen - hepatitis B virus (subtype ayw)
 N/Contains: major surface antigen; middle surface antigen
 C/Species: hepatitis B virus, HBV
 C/Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 15-Nov-1996
 C/Accession: A93217; A03703
 R/Pasek, M.; Goto, T.; Gilbert, W.; Zink, B.; Schaller, H.; Mackay, P.; Leadbetter, G.; A.
 Nature 282, 575-579, 1979
 A/Title: Hepatitis B virus genes and their expression in E. coli.
 A/Reference number: A93217; MUID:81012115; PMID:399329
 A/Accession: A93217
 A/Molecule type: DNA
 A/Residues: 1-389 <PAS>
 A/Cross-references: GB:J02202
 C/Genetic: S
 C/Superfamily: hepatitis B virus surface antigen
 C/Keywords: glycoprotein; surface antigen (gene pre-S2/S) #status predicted <DSA>
 F/109-389/Product: middle surface antigen (gene S) #status predicted <MSA>
 F/164-389/Product: major surface antigen (gene S) #status predicted
 F/4,112,166/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 96.7%; Score 59; DB 1; Length 389;
 Best Local Similarity 91.7%; Pred. No. 0.0028;
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSSIFSRIGDP 12
 DB 149 PLSSIFSRIGDP 160

RESULT 6
 SAV1AI
 Large surface antigen - hepatitis B virus (strain alpha1)
 N/Contains: major surface antigen; middle surface antigen
 C/Species: hepatitis B virus, HBV

C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Jul-1999
C/Accession: B34773; J02082
R/Tong, S.; Li, J.; Vitvitski, L.; Trepo, C.
Virology 176, 596-603, 1990
A/Title: Active hepatitis B virus replication in the presence of anti-HBs is associated
A/Reference number: A34773; MUID:90266476; PMID:234566
A/Accession: B34773
A/Status: translation not shown
A/Molecule type: DNA
A/Residues: 1-389 <TON>
A/Cross-references: EMBL:M32138; NID:9329667; PIDN:AAA45502.1; PID:9329669
R/Norder, H.; Hammes, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnus,
J. Gen. Virol. 74, 1341-1348, 1993
A/Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
A/Reference number: J02044; MUID:93329382; PMID:8336122
A/Accession: J02082
A/Molecule type: DNA
A/Residues: 164-389 <NOR>
A/Experimental source: genogroup D, subtype ayw2, strain HBV-alpha
C/Genetics:
A/Gene: pre-S1/pre-S2/S
A/Introns: 111/3
C/Superfamily: hepatitis B virus surface antigen
C/Keywords: glycoprotein, surface antigen
F/109-389/Product: middle surface antigen (gene pre-S2/S) #status predicted <DSA>
F/164-389/Product: major surface antigen (gene S) #status predicted <MSA>
F/4,112,166/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 96.7%; Score 59; DB 1; Length 389;
Best Local Similarity 91.7%; Pred. No. 0.0028;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSIFSRIGDP 12
Db 149 PLSIFSRIGDP 160

RESULT 7
S47407
surface antigen - hepatitis B virus (subtype ayw4)
N/Alternate names: envelope protein; Hbs antigen
N/Contains: surface antigen pre-S1 (large envelope protein); surface antigen pre-S2 (mid
C/Species: hepatitis B virus, HBV
A/Variety: subtype ayw4
C/Date: 23-Nov-1994 #sequence_revision 12-May-1995 #text_change 26-Aug-1999
R/Plucienczak, A.
Submitted to the EMBL Data Library, August 1994
A/Description: Molecular cloning and sequencing of two complete genomes of polish isolat
A/Reference number: S47404
A/Accession: S47407
A/Molecule type: DNA
A/Residues: 1-389 <PLU>
A/Cross-references: EMBL:Z35716; NID:9527435; PIDN:CAA84788.1; PID:9527439
A/Experimental source: subtype ayw4
R/Norder, H.; Hammes, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnus,
J. Gen. Virol. 74, 1341-1348, 1993
A/Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
A/Reference number: J02044; MUID:93329382; PMID:8336122
A/Accession: J02071
A/Molecule type: DNA
A/Residues: 164-389 <NOR>
A/Experimental source: subtype ayw3, strain 8950/90
C/Genetics:
A/Gene: S
A/Introns: 111/3
C/Superfamily: hepatitis B virus surface antigen
C/Keywords: surface antigen
F/1-389/Product: surface antigen pre-S1 (large envelope protein) #status predicted <PS1>
F/1-108/Domain: pre-S1 domain #status predicted <PRE1>
F/109-389/Product: surface antigen pre-S2 (middle envelope protein) #status predicted <F
F/109-163/Domain: pre-S2 domain #status predicted <PRE2>

F/164-389/Product: surface antigen S (small envelope protein) #status predicted <SAG>

Query Match 96.7%; Score 59; DB 2; Length 389;
Best Local Similarity 91.7%; Pred. No. 0.0028;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSIFSRIGDP 12
Db 149 PLSIFSRIGDP 160

RESULT 8
S32202
large surface antigen - hepatitis B virus (subtype ayw, isolate patient C1005)
N/Contains: major surface antigen; middle surface antigen
C/Species: hepatitis B virus, HBV
A/Variety: subtype ayw, isolate patient C1005
C/Date: 20-Feb-1995 #sequence_revision 06-Dec-1996 #text_change 06-Dec-1996
R/Prelisier-Adams, S.; Schlayer, M.J.; Peters, T.; Hettler, F.; Gerok, W.; Raenack, J.
submitted to the EMBL Data Library, March 1993
A/Description: Identification and sequence analysis of hepatitis B virus DNA in immunolog
A/Reference number: S32202
A/Accession: S32202
A/Molecule type: DNA
A/Residues: 1-389 <PRE>
A/Cross-references: EMBL:X72702
A/Experimental source: subtype ayw, isolate patient C1005
C/Genetics:
A/Gene: pre-S1/pre-S2/S
A/Introns: 111/3
C/Superfamily: hepatitis B virus surface antigen
C/Keywords: surface antigen
F/1-389/Product: large surface antigen (gene pre-S1/pre-S2/S) #status predicted <DS1>
F/109-389/Product: middle surface antigen (gene pre-S2/S) #status predicted <DSA>
F/164-389/Product: major surface antigen (gene S) #status predicted <MSA>

Query Match 96.7%; Score 59; DB 2; Length 389;
Best Local Similarity 91.7%; Pred. No. 0.0028;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSIFSRIGDP 12
Db 149 PLSIFSRIGDP 160

RESULT 9
T13474
large surface antigen - hepatitis B virus (isolate 29Y11HCC)
N/Alternate names: envelope protein; Hbs antigen
N/Contains: major surface antigen; middle surface antigen
C/Species: hepatitis B virus, HBV
A/Variety: isolate 29Y11HCC
C/Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 21-Jul-2000
R/Takahashi, K.; Akahane, Y.; Hino, K.; Ohta, Y.; Mashiro, S.
Arch. Virol. 143, 2313-2326, 1998
A/Title: Hepatitis B virus genomic sequence in the circulation of hepatocellular carcinoma
A/Reference number: Z17684; MUID:99129050; PMID:9930189
A/Accession: T13474
A/Molecule type: DNA
A/Residues: 1-382 <TRK>
A/Cross-references: EMBL:AB014388; NID:93582381; PIDN:BAV3936.1; PID:93582384
A/Experimental source: Japanese patient with hepatocellular carcinoma isolate 29Y11HCC
C/Genetics:
A/Gene: S
A/Introns: 123/2
C/Superfamily: hepatitis B virus surface antigen
C/Keywords: glycoprotein, surface antigen

Query Match 88.5%; Score 54; DB 2; Length 382;
Best Local Similarity 83.3%; Pred. No. 0.025; Indels 0; Gaps 0;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 PLSIFSRIGDP 12
 |||||
 Db 142 PLSIFSRIGDP 153

RESULT 10

113469
 large surface antigen - hepatitis B virus (isolate 09D09HCC)
 N/Alternate names: envelope protein
 N/Contains: major surface antigen; middle surface antigen
 C/Species: hepatitis B virus, HBV
 C/Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 21-Jul-2000
 C/Accession: J13469
 R/Takahashi, K.; Akahane, Y.; Hino, K.; Ohta, Y.; Mishiro, S.
 Arch. Virol. 143, 2313-2326, 1998
 A/Title: Hepatitis B virus genomic sequence in the circulation of hepatocellular carcin
 A/Reference number: Z17684; MUID:99129050; PMID:9930189
 A/Accession: J13469
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-384 <TAK>
 A/Cross-references: EMBL:AB014368; NID:93551304; PIDN:BA032864.1; PID:G3551308
 A/Experimental source: Japanese patient with hepatocellular carcinoma isolate 09D09HCC
 C/Genetics:
 A/Gene: S
 A/Intron: 123/2
 C/Superfamily: hepatitis B virus surface antigen
 C/Keywords: glycoprotein; surface antigen

Query Match 88.5%; Score 54; DB 2; Length 384;
 Best Local Similarity 83.3%; Pred. No. 0.025; Indels 0; Gaps 0;
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 PLSIFSRIGDP 12
 |||||
 Db 144 PLSIFSRIGDP 155

RESULT 11

620745
 surface antigen - hepatitis B virus (subtype ayw, patient C)
 N/Alternate names: HBs antigen
 N/Contains: surface antigen pre-S1 (large envelope protein); surface antigen pre-S2 (mid
 C/Species: hepatitis B virus, HBV
 A/Variety: subtype ayw, patient C
 C/Date: 20-Feb-1995 #sequence_revision 01-Nov-1996 #text_change 26-Aug-1999
 C/Accession: S20745
 R/Jual, M.B.; Mazzoleni, A.P.; Balesier, A.; Melle, A.; Porru, A.
 submitted to the EMBL Data Library, March 1992
 A/Description: Sequence analysis of HBV genomes isolated from patients with HBsAg negativ
 A/Reference number: S20745
 A/Accession: S20745
 A/Molecule type: DNA
 A/Residues: 1-388 <LAI>
 A/Cross-references: EMBL:X65257; NID:G59429; PIDN:CAA46349.1; PID:G59430
 A/Experimental source: subtype ayw, patient C
 C/Genetics:
 A/Gene: S
 C/Superfamily: hepatitis B virus surface antigen
 C/Keywords: surface antigen
 F/1-389/Product: surface antigen pre-S1 (large envelope protein) #status predicted <PS1>
 F/1-108/Dominant: pre-S1 domain #status predicted <PRE1>
 F/109-389/Product: surface antigen pre-S2 (middle envelope protein) #status predicted <P
 F/109-163/Dominant: pre-S2 domain #status predicted <PRE2>
 F/164-389/Product: surface antigen S (small envelope protein) #status predicted <PSD>

Query Match 88.5%; Score 54; DB 2; Length 389;
 Best Local Similarity 100.0%; Pred. No. 0.025; Indels 0; Gaps 0;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 PLSIFSRIGDP 12

Db 150 PLSIFSRIGDP 160
 |||||

RESULT 12

SAVLA
 large surface antigen - hepatitis B virus (subtype adr)
 N/Contains: major surface antigen; middle surface antigen
 C/Species: hepatitis B virus, HBV
 C/Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 07-May-1999
 C/Accession: A03705; S04569; JQ2107; P00608
 R/Ono, Y.; Ono, H.; Sasaki, R.; Igatahshi, K.; Sugino, Y.; Nishioka, K.
 Nucleic Acids Res. 11, 1747-1757, 1983
 A/Title: The complete nucleotide sequences of the cloned hepatitis B virus DNA; subtype
 A/Reference number: A93460; MUID:83168919; PMID:630076
 A/Accession: A03705
 A/Molecule type: DNA
 A/Residues: 1-400 <ONO>
 A/Cross-references: GB:V00867
 R/Rho, H.M.; Kim, K.; Hyun, S.W.; Kim, Y.S.
 Nucleic Acids Res. 17, 2124, 1989
 A/Title: The nucleotide sequence and reading frames of a mutant hepatitis B virus subtype
 A/Reference number: S04568; MUID:89183619; PMID:2928116
 A/Accession: S04569
 A/Status: translation not shown

A/Molecule type: DNA
 A/Residues: 1-50, 'R', 52-66, 'YP', 69-129, 'V', 131-142, 'P', 144-164, 'S', 166-176, 'S', 178-338, 'C'
 A/Cross-references: EMBL:X14193
 R/Norder, H.; Hammam, B.; Lee, S.D.; Bile, K.; Courouge, A.M.; Muehahwar, I.K.; Magnus,
 J. Gen. Virol. 74, 1341-1348, 1993
 A/Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
 A/Reference number: JQ2044; MUID:93329382; PMID:8336122
 A/Accession: JQ2107
 A/Molecule type: DNA
 A/Residues: 175-400 <NOR>

A/Experimental source: genogroup C, subtype adr, strain PBRBadr4
 R/Norder, H.; Courouge, A.M.; Magnus, L.O.

J. Gen. Virol. 73, 3141-3145, 1992

A/Title: Molecular basis of hepatitis B virus serotype variations within the four major

A/Reference number: P00453; MUID:93107848; PMID:1469353

A/Accession: P00608

A/Molecule type: DNA

A/Residues: 275-354 <NO2>

A/Experimental source: subtype adrq+, Bau

C/Genetics:

A/Gene: pre-S1/pre-S2/S

C/Superfamily: hepatitis B virus surface antigen

C/Keywords: glycoprotein; surface antigen (gene pre-S2/S) #status predicted <DSA>

F/120-400/Product: middle surface antigen (gene S) #status predicted <MSA>

F/175-400/Product: major surface antigen (gene S) #status predicted <MSA>

F/15,123,177/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 88.5%; Score 54; DB 1; Length 400;
 Best Local Similarity 83.3%; Pred. No. 0.026; Indels 0; Gaps 0;
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 PLSIFSRIGDP 12
 |||||
 Db 160 PLSIFSRIGDP 171

RESULT 13

S35528
 surface antigen - hepatitis B virus (subtype adr)
 N/Alternate names: envelope protein; HBs antigen
 N/Contains: surface antigen pre-S1 (large envelope protein); surface antigen pre-S2 (midc
 C/Species: hepatitis B virus, HBV
 A/Variety: subtype adr
 C/Date: 09-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 20-Jun-2000
 C/Accession: S35528
 R/Mukaida, M.; Kumazawa, T.; Hoshi, A.; Kawaguchi, R.; Hiki, J. K.
 Nucleic Acids Res. 20, 6105, 1992
 A/Title: The complete nucleotide sequence of hepatitis B virus, subtype adr (SRAD) and

A:Reference number: S35527; MUID:93096607; PMID:1461746
A:Accession: S35528
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-400 <MUT>
A:Cross-references: EMBL:DJ2980; NID:g221500; PIDD:BAA0359.1; PID:g221502
A:Experimental source: subtype adr
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1992
C:Genetics:
A:Gene: S
A:Introns: 165/2
C:Superfamily: hepatitis B virus surface antigen
C:Keywords: surface antigen
F:1-400/Product: surface antigen pre-S1 (large envelope protein) #status predicted <PS1>
F:1-119/Domin: pre-S1 domain #status predicted <PRE1>
F:120-400/Product: surface antigen pre-S2 (middle envelope protein) #status predicted <PRE2>
F:120-174/Domin: pre-S2 domain #status predicted <PRE2>
F:175-400/Product: surface antigen S (small envelope protein) #status predicted <PSD>

Query Match 88.5%; Score 54; DB 2; Length 400;
Best Local Similarity 83.3%; Pred. No. 0.026;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLSIFSRIGDP 12
DB 160 PLSIFSRIGDP 171

RESULT 14
S34392
surface antigen - hepatitis B virus (subtype adr)
N:Alternate names: Hbs antigen
M:Contains: surface antigen pre-S1 (large envelope protein); surface antigen pre-S2 (middle envelope protein); subtype adr
C:Species: hepatitis B virus, HBV
A:Variety: subtype adr
C:Date: 07-Sep-1994 #sequence revision 10-Nov-1995 #text_change 26-Aug-1999
C:Accession: S34392; PNO602
R:Loncarevic, I.F.; Zentgraf, H.; Schroeder, C.H.
Nucleic Acids Res. 18, 4340, 1990
A:Title: Sequence of a replication competent hepatitis B virus genome with a prex open r
A:Reference number: S12598; MUID:90370503; PMID:2395664
A:Accession: S34392
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-445 <LON>
A:Cross-references: EMBL:X52939; NID:9457780; PIDD:CAA37114.1; PID:9457783
A:Experimental source: subtype adr
R:Mimm, L.T.; Solomon, L.R.; Ebert, J.W.; Fields, H.
Biochem. Biophys. Res. Commun. 195, 186-191, 1993
A:Title: Unique Pres sequence in a gibbon-derived hepatitis B virus variant.
A:Reference number: PNO601; MUID:93371402; PMID:8363598
A:Accession: PNO601
A:Molecule type: protein
A:Residues: 57-95 <MIM>
A:Accession: PNO602
A:Molecule type: protein
A:Residues: 155-196 <MI2>
C:Genetics:
A:Gene: S
A:Introns: 210/3
C:Superfamily: hepatitis B virus surface antigen
C:Keywords: surface antigen
F:1-445/Product: surface antigen pre-S1 (large envelope protein) #status predicted <PS1>
F:1-164/Domin: pre-S1 domain #status predicted <PRE1>
F:165-445/Product: surface antigen pre-S2 (middle envelope protein) #status predicted <PRE2>
F:165-219/Domin: pre-S2 domain #status predicted <PRE2>
F:220-445/Product: surface antigen S (small envelope protein) #status predicted <PSD>

DB 205 PLSIFSRIGDP 216

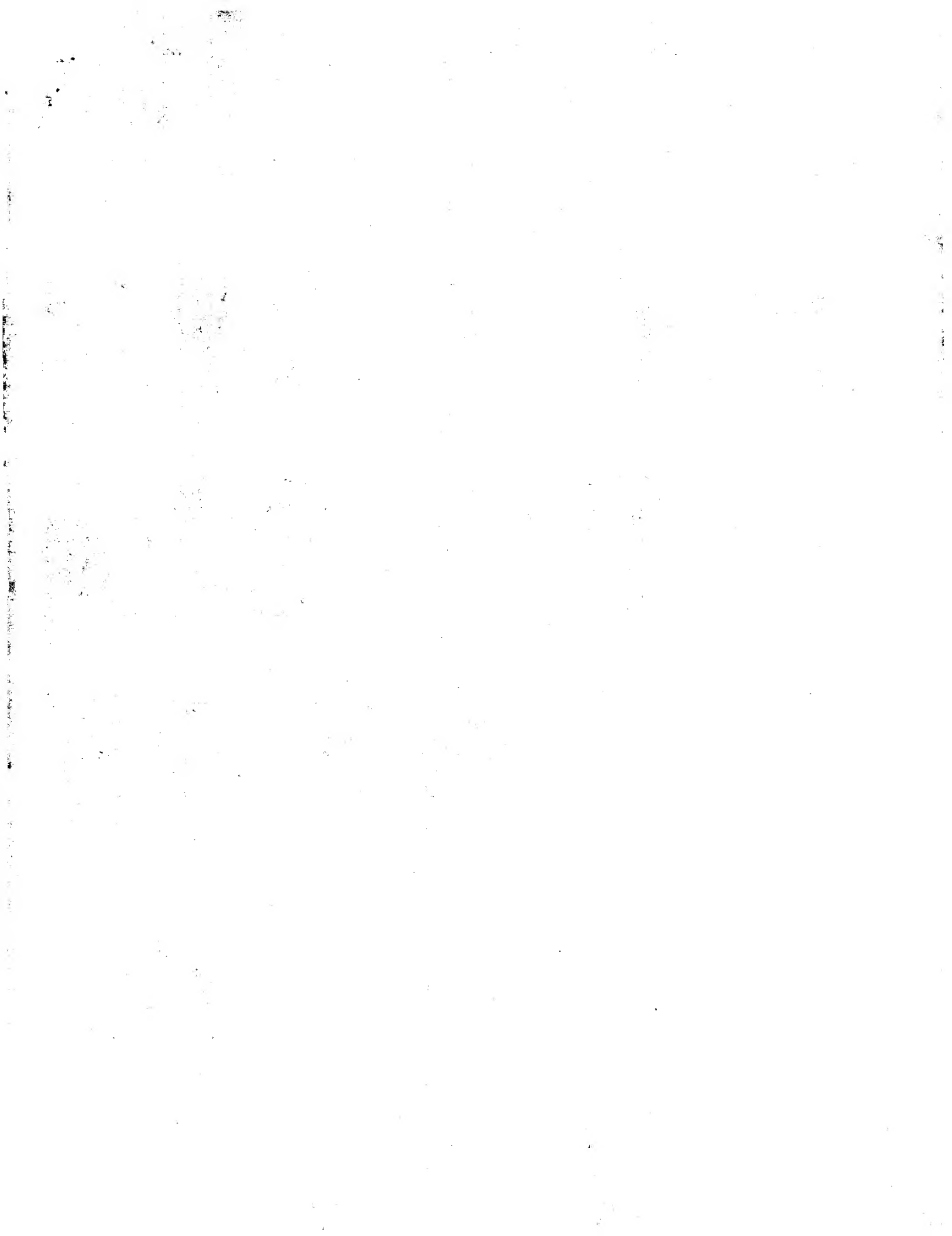
RESULT 15
S67506
large surface antigen - hepatitis B virus (subtype ayw3, isolate Hope CH1357)
N:Alternate names: envelope protein, Hbs antigen
M:Contains: surface antigen pre-S1 (large envelope protein); surface antigen pre-S2 (middle envelope protein); subtype ayw3, isolate Hope CH1357
C:Species: hepatitis B virus, HBV
A:Variety: subtype ayw3, isolate Hope CH1357
C:Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 26-Aug-1999
C:Accession: S67506
R:Norde, H.; Ebert, J.W.; Fields, H.A.; Mushahwar, I.K.; Magnius, L.O.
Virology 218, 214-223, 1996
A:Title: Complete sequencing of a gibbon hepatitis B virus genome reveals a unique genov
A:Reference number: S67503; MUID:96207410; PMID:8615024
A:Accession: S67506
A:Molecule type: DNA
A:Residues: 1-389 <NOR>
A:Cross-references: EMBL:U46935; NID:91814218; PIDD:AB41952.1; PID:91814222
A:Experimental source: subtype ayw3, isolate Hope CH1357
C:Genetics:
A:Gene: pre-S1/pre-S2/S
A:Introns: 112/2
C:Superfamily: hepatitis B virus surface antigen
C:Keywords: glycoprotein, surface antigen
F:1-389/Product: surface antigen pre-S1 (large envelope protein) #status predicted <PS1>
F:1-108/Domin: pre-S1 specific domain #status predicted <PRE1>
F:109-389/Product: surface antigen pre-S2 (middle envelope protein) #status predicted <PRE2>
F:109-163/Domin: pre-S2 specific domain #status predicted <PRE2>
F:164-389/Product: surface antigen S (small envelope protein) #status predicted <PSD>
F:4,112,166/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 77.0%; Score 47; DB 2; Length 389;
Best Local Similarity 81.8%; Pred. No. 0.54;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PLSIFSRIGDP 12
DB 150 PLSIFSRIGDP 160

Search completed: February 3, 2003, 09:26:29
Job time: 1.79443 secs

QY 1 PLSIFSRIGDP 12



PT of viral protein, permeability peptide and cell-binding site -
 XX
 PS Disclosure, Page 2; 34pp; German.
 CC This invention describes a novel particle (A), comprising a protein coat
 CC with a fusion protein (FP), and, inside the coat, a nucleic acid (I)
 CC including the sequence for a virus-specific packaging signal (psi) and a
 CC structural gene. FP contains a virus protein (VP), a peptide (P) that
 CC mediates cell permeability and a heterologous cell-specific binding site
 CC (RCD). The invention also describes (1) producing (A) in which FP
 CC contains an LHBs (large surface protein of hepatitis B virus (HBV)) and
 CC RGD; (2) preparing (A) in which FP contains an HBV core antigen (HBcAg),
 CC (P) and RGD; (3) FP; (4) DNA encoding FP; and (5) expression vector
 CC containing the DNA of (d). The products of the invention are used in gene
 CC therapy of cells and tissues, in vivo or ex vivo. This sequence
 CC represents a fragment of the Hepatitis B virus cell permeability peptide
 CC which is described in the method of the invention.
 CC
 SQ Sequence 12 AA;
 Query Match 100.0%; Score 61; DB 21; Length 12;
 Best Local Similarity 100.0%; Pred. No. 3.7e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 PLSSIFSRIGDP 12
 DB 1 PLSSIFSRIGDP 12
 RESULT 2
 AA87906
 ID AA87906 standard; Protein; 12 AA.
 XX
 AC AA87906;
 XX
 DT 01-SEP-2000 (first entry)
 XX
 DE HBV cell surface protein ZPP fragment.
 XX
 KW Cell surface protein; ZPP; cell permeability mediating polypeptide;
 XX therapy; treatment.
 OS Hepatitis B virus.
 XX
 PN DE19850718-C1.
 XX
 PD 18-MAY-2000.
 XX
 PF 03-NOV-1998; 98DE-1050718.
 XX
 PR 03-NOV-1998; 98DE-1050718.
 XX
 PA (HILD/) HILDT E.
 XX
 PI HILDT E, Oess S;
 XX
 DR WPI; 2000-340689/30.
 DR N-PSDB; AAA39621.
 XX
 PT Novel cell permeability-mediating polypeptide useful for mediating
 PT permeability of substances such as other polypeptides in cells -
 XX
 PS Claim 1; Fig 1; 8pp; German.
 XX
 CC This invention describes a novel cell permeability mediating polypeptide
 CC (CPM) comprising (A) or a sequence which differs by one or more amino
 CC acids, which shows cell permeability and is not hepatitis B virus (HBV)
 CC surface protein and where the DNA sequence of the latter amino acid
 CC sequence hybridizes with the DNA of (I). CPM can be used to mediate cell
 CC permeability of substances, especially polypeptides, nucleic acids and
 CC chemical connections. This is useful for increasing the reactivity of
 CC such substances in therapeutic treatment of various conditions. This
 CC sequence represents a fragment of the HBV cell surface protein ZPP which

CC is described in the method of the invention.
 XX
 SQ Sequence 12 AA;
 Query Match 100.0%; Score 61; DB 21; Length 12;
 Best Local Similarity 100.0%; Pred. No. 3.7e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 PLSSIFSRIGDP 12
 DB 1 PLSSIFSRIGDP 12
 RESULT 3
 AAR14747
 ID AAR14747 standard; peptide; 25 AA.
 XX
 AC AAR14747;
 XX
 DT 28-JAN-1992 (first entry)
 XX
 DE HBV Pre S2 peptide.
 XX
 KW Monoclonal antibodies; hepatitis B virus; envelope; vaccine;
 KW epitope; antigen.
 XX
 OS Synthetic.
 XX
 PN EP456215-A.
 XX
 PD 13-NOV-1991.
 XX
 PF 08-MAY-1991; 91EP-0107488.
 XX
 PR 11-MAY-1990; 90US-0522286.
 XX
 PA (ABBO) ABBOTT LABORATORIES.
 XX
 PI MImms LT, Floreani MF,
 XX
 DR WPI; 1991-334198/46.
 XX
 PT New MAb to PreS2 and PreS1 polypeptide(s) of Hepatitis-B
 PT envelope - used to sub-type HBV in sample and as vaccine against
 PT HBV.
 XX
 ES Claim 1; Page 22; 24pp; English.
 XX
 CC The peptide corresponds to residues 150-174 of the HBV PreS2 protein.
 CC It was used to raise monoclonal antibodies specific for the HBV M
 CC protein but which do not bind to the HBV S or L proteins. The MAb
 CC can be used to sub-type HBV and to prepare subunit vaccines for HBV.
 XX
 SQ Sequence 25 AA;
 Query Match 100.0%; Score 61; DB 12; Length 25;
 Best Local Similarity 100.0%; Pred. No. 8.3e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 PLSSIFSRIGDP 12
 DB 11 PLSSIFSRIGDP 22
 RESULT 4
 AAR17023
 ID AAR17023 standard; peptide; 36 AA.
 XX
 AC AAR17023;
 XX
 DT 18-APR-2002 (first entry)
 XX
 DE Hepatitis B virus (HBV) ayw subtype pre-S2 region.

XX	Hepatitis B virus; HBV, core antigen; HBeAg; immune system; typhoid;
KM	prophylactic; gene therapy; vaccine; hepatitis A virus; HAV; herpes;
KM	hepatitis C virus; HCV; influenza; foot-and-mouth disease; diarrhoea;
KW	tuberculosis; polio; rabies; acquired immunodeficiency syndrome; AIDS;
KM	dengue fever; yellow fever; malaria; whooping cough; salmonellosis;
KM	food poisoning; meningitis; gonorrhoea; antiviral; antibacterial;
XX	antiprotocozal; pre-S2 region.
OS	Hepatitis B virus.
PN	WO200198333-A2.
XX	
PD	27-DEC-2001.
XX	
PF	22-JUN-2001; 2001MO-GB02817.
XX	
PR	22-JUN-2000; 2000GB-0015308.
XX	
PR	06-OCT-2000; 2000GB-0024544.
PA	(CELL-) CELLTECH PHARM LTD.
XX	
PI	Page M, Li J, Pumpens P;
XX	
DR	WPI; 2002-098223/13.
XX	
N-PSDB;	AAD27424.
XX	
PT	New proteins comprising a modified hepatitis B core antigen, useful as
PT	a vaccine in prophylactic or therapeutic vaccination of the human or
XX	animal body, particularly against hepatitis B virus infection -
XX	Disclosure, Fig 3A; 40pp; English.
XX	
SS	The invention relates to modified proteins comprising hepatitis B virus
CC	(HBV) core antigen (HBeAg) wherein one or more of the four arginine
CC	repeats has been deleted and the protein comprising the C-terminal
CC	cysteine of HBeAg. The deleted region may be replaced by an epitope
CC	from a protein other than HBeAg, in which case the HBeAg acts as a
CC	carrier to present the epitope to the immune system. This chimeric
CC	protein or its nucleic acid is useful as a vaccine or in a method of
CC	prophylactic or therapeutic vaccination of the human or animal body,
CC	particularly against HBV. The nucleic acid encoding the protein may
CC	be used in gene therapy or DNA vaccination protocols. The chimeric
CC	protein or its nucleic acid may also be used as the basis of a
CC	prophylactic vaccine against a range of diseases, e.g. HBV, hepatitis
CC	C virus (HAV), hepatitis C virus (HCV), influenza, foot-and-mouth
CC	disease, polio, herpes, rabies, acquired immunodeficiency syndrome
CC	(AIDS), dengue fever, yellow fever, malaria, tuberculosis, whooping
CC	cough, salmonellosis, typhoid, food poisoning, diarrhoea, meningitis
CC	or gonorrhoea. The present sequence is Hepatitis B virus (HBV) ayw
CC	subtype pre-S2 region.
XX	
SQ	Sequence 36 AA;
XX	
Query Match	100.0%; Score 61; DB 23; Length 36;
Best Local Similarity	100.0%; Pred. No. 0.00012;
Matches 12; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Dy	1 PLSSIFSRIQDP 12
Db	22 PLSSIFSRIQDP 33
RESULT 5	
ID	AAR76747 standard; peptide; 60 AA.
AA	AAR76747
AC	
XX	AAR76747;
DT	18-MAR-1996 (first entry)
XX	
DE	HBV surface antigen pre-S2 region and FlmH residues 224-226.
XX	

KM	FtmH; type I fimbriae; organelle; adhesin; alpha-D-mannoside residue;
KW	FtmA; FtmF; FtmG; receptor binding site; PCR; amplity.
XX	
OS	Chimeric - Hepatitis B virus.
XX	Chimeric - Escherichia coli.
PH	
Key	Location/Qualifiers
FT	Misc-difference 1
PT	/note= "Represents FtmH residue 224"
FT	Peptide
PT	2...4
FT	/note= "Linker peptide"
PT	Peptide
FT	5...56
PT	/note= "HBV surface antigen pre-S2 region"
FT	Peptide
PT	57..59
FT	/note= "Linker peptide"
PT	Misc-difference 60
PN	/note= "Represents FtmH residue 226"
XX	
MO9520657-AI.	
XX	
03-AUG-1995.	
PD	
XX	
PF	27-JAN-1995; 95WO-DK00042.
XX	
PR	27-JAN-1994; 94US-0187166.
XX	
PA	(GXBI-) GX BIOSYSTEMS AS.
PI	Hasty DL, Klemm P, Molin S, Pallesen L, Sokurenko EV;
PI	WPI, 1995-275442/36.
DR	N-PSDB; AAQ93059.
XX	
Receptor specific bacterial adhesins - useful for targeting active	
compounds and microbial cells to locations of receptors	
Example 3; Page 57; 152pp; English.	
PS	
XX	This sequence is encoded by a fragment of the plasmid pLP3A8, which
CC	comprises a ftmH fusion gene comprising the sequence encoding the pre-S2
CC	region of the hepatitis B viral surface antigen inserted into the ftmH
CC	gene. The primer sequences given in AAQ93057-58 caused the insertion of
CC	the pre-S2 region into the ftmH protein at position 224-226. The
CC	resulting chimeric gene was then further modified by insertion of the
CC	cholera toxin B chain into a different position of the ftmH adhesin of
CC	type 1 fimbriae. Restriction site handles (BglII-sites) were introduced
CC	into the ftmH gene, and the foreign epitopes are then inserted in-frame.
CC	In the selected positions the insertion of the epitopes did not
CC	significantly alter the adhesive function of the ftmH protein. The
CC	expression of the chimeric proteins on the surface of fimbriae on
CC	as bacterial hosts illustrated the possibility of using bacterial adhesins
CC	as general presenters of foreign antigens and epitopes. These chimeric
CC	genes may be used in the production of variant ftmH adhesins which may
CC	be useful for targeting active compounds and microbial cells to
CC	locations comprising selected receptors to which the adhesins bind.
XX	
SQ	Sequence 60 AA;
Query Match	100.0%; Score 61; DB 16; Length 60;
Best Local Similarity	100.0%; Pred. No. 0.00021;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	1 PLSSIFSRIGDP 12
Db	42 PLSSIFSRIGDP 53
RESULT 6	
ID	AAAY30947 standard; Protein; 102 AA.
XX	
AC	AAAY30947;
XX	

```

XX      20-Oct-1999   (first entry)
DE      Human HbsAg (Pres2 plus S region) truncated peptide.
XX
XX      Vaccine; accessory molecule; antigen; transport; presentation; allergy;
KM      antibacterial; antiviral; antifungal; anti-allergic; antidiabetic;
KM      anti-inflammatory; anti-arthritis; anti-asthma; anticancer; treatment;
KM      prevention; infection; bacterial; viral; fungal; autoimmune disease;
KM      rheumatoid arthritis; diabetes; multiple sclerosis; pancreatitis;
KM      inflammatory condition; psoriasis; immune deficiency; metastases;
KM      cancer; Th1-type helper; response; humoral; cellular response; B cell;
KM      T cell; proliferation; immunoglobulin synthesis; isotype switching;
KM      immunomodulator; immune response; asthma; human; HbsAg.
XX
OS      Homo sapiens.
XX
FN      W09941368-A2.
XX
PD      19-AUG-1999.
XX
PF      10-FEB-1999;    99WO-US03020.
XX
PR      11-FEB-1998;    98US-0074294.
XX      11-FEB-1998;    98US-0021769.
PA      (MAXY-) MAXYGEN INC.
PI      Howard R, Punnonen J, Stemmer WPC, Whalen RG;
PT      WPI, 1999-508645/42.
XX
PT      Identifying nucleic acid that directly or indirectly modulates the
XX      immune response to a genetic vaccine vector, e.g. for prevention of
XX      infection or cancer
XX
PS      Example 3; Page 70; 105pp; English.
XX
CC      This invention describes the identification of a novel polynucleotide (I)
CC      that modulates the immune response to a genetic vaccine vector (A), or
CC      encodes a polypeptide (II) with similar effect. The invention also
CC      describes the identification of a polynucleotide (Ia) encoding an
CC      accessory molecule (Iia) that improves transport and presentation of
CC      antigen by a cell. The products of the invention have antibacterial,
CC      antiviral, antifungal, anti-allergic, antidiabetic, anti-inflammatory,
CC      anti-arthritis, anti-asthma, anticancer and immunomodulatory. Optimized
CC      (I) are incorporated into (A), or (I) or its encoded (II) are
CC      administered together with (A). (A) are used to treat or prevent
CC      infections (bacterial, viral or fungal), autoimmune disease (e.g.
CC      rheumatoid arthritis, diabetes or multiple sclerosis), other inflammatory
CC      conditions (e.g. psoriasis or pancreatitis), immune deficiency, allergy,
CC      asthma or cancer (including metastases). (I) are also used for
CC      recombinant production of (II). (I) make it possible to tailor an immune
CC      response to particular requirements, e.g. to direct a Th1-type helper
CC      response to increase humoral or cellular responses (functioning as
CC      adjuvant) to control B or T cell proliferation to induce immunoglobulin
CC      synthesis or isotype switching. This sequence represents the HbsAg
CC      (Pres2 plus S region) truncated peptide which is used in the method of
CC      the invention.
XX
SQ      Sequence      102 AA;
XX
Query Match          100.0%; Score 61; DB 20; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.00038;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
OY      1 PLSISFRIQDP 12
        |||||
Db       41 PLSISFRIQDP 52

```

XX	AA93740;
AC	03-OCT-2000 (first entry)
DT	Amino acid sequence for pre-S peptide.
DE	Plant expression vector; antigen; hepatitis B surface antigen; HBsAg;
XX	vaccine; ss.
XX	Synthetic.
OS	WO200037610-A2.
EN	29-JUN-2000.
PD	23-DEC-1999; 99WO-US31020.
XX	23-DEC-1998; 98US-0113827.
PR	(BOYC-) BOYCE THOMPSON INST PLANT RES.
PA	(HEAL-) HEALTH RES INST.
XX	Mason HS, Thanavala Y, Amren CJ, Richter E;
PI	wpi; 2000-452181/39.
XX	N-Psdb; AAA47000.
DR	New expression vector for transforming plants comprising two expression
PT	cassettes useful for producing plant material comprising anti-hepatitis
PT	B antibodies -
PS	Disclosure; Fig 5; 14dp; English.
XX	The specification describes a plant expression vector which comprises
CC	two expression cassettes, the first comprising a polynucleotide encoding
CC	an antigen and the second comprising a non-identical polynucleotide
CC	encoding the same antigen. The antigen is especially a hepatitis B
CC	surface antigen (HBsAg). The expression vector is used to transform
CC	bacterial and plant cells to elicit the production of anti-hepatitis B
CC	antibodies and are useful as vaccines. The present sequence represents
CC	a pre-S peptide.
XX	
SQ	Sequence 164 AA;
Query Match	100.0%; Score 61; DB 21; Length 164;
Best Local Similarity	100.0%; Pred. No. 0.00064;
Matches 12; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Oy	1 PLSIFSRIQDP 12
Db	149 PLSIFSRIQDP 160
RESULT 8	
ID	AA90760
AY	AA90760 standard; Protein; 164 AA.
XX	AA90760;
DT	18-AUG-2000 (first entry)
XX	Hepatitis B virus pres protein sequence.
DE	Nuclear localisation signal; NLS; SV40; large T cell antigen;
KM	fusion protein; diagnosis; detection; variant; vaccine; chaperone;
KM	coprecipitate.
XX	Hepatitis B virus.
OS	WO200020606-A1.
EN	13-APR-2000.
PD	

XX 02-OCT-1998; 98MO-EP06298.
 PF 02-OCT-1998; 98MO-EP06298.
 XX (REIM/) REIMANN H.
 PA (SCHL/) SCHIRMbeck R.
 XX Reimann H, Schirmbeck R;
 PI WPI; 2000-328936/28.
 DR Novel polynucleotide encoding a fusion protein which is stable in a
 PT cell, useful for the production of peptides which coprecipitate
 PT chaperones, by using truncated variants of the SV40 large T antigen
 XX with an intact N-terminus -
 XX Example 1; Page 26; 57pp; English.
 PS The present invention describes a polynucleotide (1) encoding a fusion
 CC protein (FP) (11) which is stable in a cell, and comprises a first
 CC polypeptide (P1) and a second polypeptide (P2) which co-precipitates a
 CC chaperone. The invention also provides methods for the production of
 CC fusion proteins which comprise epitopes. Compositions comprising the
 CC fusion proteins or epitopes of the invention are used as vaccines, and
 CC for the production of antibodies. The methods provide an inexpensive
 CC and efficient means of reliably expressing (poly)peptides comprising
 CC epitopes. The present sequence represents the Hepatitis B virus (HBV)
 CC pres protein sequence, which is used in an example from the present
 CC invention.
 XX SQ Sequence 164 AA;
 XX
 XX Query Match 100.0%; Score 61; DB 21; Length 164;
 XX Best Local Similarity 100.0%; Pred. No. 0.00064;
 XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PLSIFSRIGDP 12
 DB 150 PLSIFSRIGDP 161
 XX
 XX RESULT 9
 XX AAR15617
 XX ID AAR15617 standard; protein; 174 AA.
 XX AAR15617;
 XX 25-MAR-1992 (first entry)
 XX DB HBSAg pre-S region subtype ayw.
 XX T-cell epitope; vaccine; hepatitis B virus; antigen.
 XX Synthetic.
 XX WO9117768-A.
 XX 28-NOV-1991.
 XX 10-MAY-1991; 91WO-US03268.
 XX 11-MAY-1990; 90US-0522663.
 XX (SCRI) SCRIPPS CLINIC & RE.
 XX MIllich DR, Thornton GB;
 XX WPI; 1991-369007/50.
 XX Hepatitis B virus surface antigen epitope(s) - useful as vaccines,
 XX immunogens or diagnostic reagents

PS Claim 1; Fig 1; 91pp; English.
 XX The amino acid sequence is that of a pre-S T cell epitope polypeptide
 CC of the pre-S (2) region of hepatitis B surface antigen Y (HBSAg/Y).
 CC It can be used to prime or vaccinate a host to induce responsiveness
 CC to HBV vaccine. The T cell epitope polypeptides can also be used as
 CC immunogens that prime T cells that respond to native HBSAg B cell
 CC epitope polypeptide. The T cell epitope polypeptides are also useful
 CC as substitutes for carrier immunogens such as KIH and are safe,
 CC defined and T cell-active. In addition to their use as vaccines, the
 CC polypeptides can be used as immunogens for prodn. of antibodies. See
 XX also AAR15618-R15622.
 XX SQ Sequence 174 AA;
 XX
 XX Query Match 100.0%; Score 61; DB 12; Length 174;
 XX Best Local Similarity 100.0%; Pred. No. 0.00068;
 XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PLSIFSRIGDP 12
 DB 160 PLSIFSRIGDP 171
 XX
 XX RESULT 10
 XX AAR23867
 XX ID AAR23867 standard; protein; 174 AA.
 XX AAR23867;
 XX 11-NOV-1992 (first entry)
 XX DE Pre-S gene region translation product (1).
 XX env gene; hepatitis B virus; HBV; S-protein; hydrophilic;
 XX hepadnavirus; immune response; antibody; envelope.
 XX Hepatitis B virus, subtype ayw.
 XX EP485361-A.
 XX 13-MAY-1992.
 XX 25-APR-1987; 87EP-0100663.
 XX 28-APR-1986; 86US-0856522.
 XX (CALY) CALIFORNIA INST TECHNOLOGY.
 XX (NYBL-) NY BLOOD CENT INC.
 XX Kent SBH, Neurath AR,
 XX WPI; 1992-161100/20.
 XX Hepatitis B pre-S peptide immunogen and vaccine - for treatment
 XX and diagnosis of hepatitis B
 XX Disclosure; Fig 2; 59pp; English.
 XX The sequences given in AAR23867 - AAR23871 are amino acid sequences
 CC deduced from sequences of the pre-S portion of the env genes
 CC corresponding to several hepatitis B virus (HBV) subtypes. These
 CC proteins have properties distinct from those of the S-protein in
 CC that they have high hydrophilicity and a high percentage of charged
 CC residues, an absence of Cys residues, the highest subtype-dependant
 CC variability among HBV DNA gene products and little homology with
 CC analogous sequences corresponding to nonhuman hepadnavirus. These
 CC properties suggest that the pre-S gene coded portion of the HBV
 CC envelope is exposed to the surface of the virion, is a target for
 CC the host's immune response and is responsible for the host range of HBV
 CC (limited to humans and some primates). Synthetic peptides and
 CC antibodies against them, having predetermined specificity offer the
 CC opportunity to explore the biological role of the pre-S protein moiety

CC of the HBV envelope. Portions of these amino acid sequences can be
 CC used in a vaccine or in diagnostics for the detection of antigens and
 CC antibodies, esp. those for the pre-S gene in sera of HBV infected
 CC humans and certain animals, eg. chimpanzees.

XX Sequence 174 AA;

Query Match 100.0%; Score 61; DB 13; Length 174;

Best Local Similarity 100.0%; Pred. No. 0.00068; Mismatches 0; Indels 0; Gaps 0;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 PLSISFSTRIGDP 12

160 PLSISFSTRIGDP 171

RESULT 11

AAB10597 AAB10597 standard; Protein; 215 AA.

AC AAB10597;

XX 08-JAN-2001 (first entry)

DE HBV fusion protein comprising HBcAg and RGD.

XX Fusion protein; protein coat; virus-specific packaging signal; psi;

KM virus protein; cell permeability; cell-specific binding site; LHB;

KM large surface protein; core antigen; gene therapy.

OS Hepatitis b virus.

OS Synthetic.

PN WO200046376-A2.

XX 10-AUG-2000.

XX 04-FEB-2000; 2000WO-DE00363.

XX 05-FEB-1999; 99DE-1004800.

XX (HILDT) HILDT E.

PA Hilde E, Hofschneider P;

XX WPI; 2000-514959/46.

DR N-PSDB; AAA71735.

XX Particle for cell-specific gene delivery, useful in gene therapy;

PT comprises nucleic acid in protein coat that includes a fusion protein

PT of viral protein, permeability peptide and cell-binding site -

XX Claim 14; Fig 2; 34pp; German.

XX This invention describes a novel particle (A), comprising a protein coat

CC with a fusion protein (FP), and, inside the coat, a nucleic acid (I)

CC including the sequence for a virus-specific packaging signal (psi) and a

CC structural gene. FP contains a virus protein (VP), a peptide (P) that

CC mediates cell permeability and a heterologous cell-specific binding site

CC (RBD). The invention also describes (1) producing (A) in which FP

CC contains an LHBs (large surface protein of hepatitis B virus (HBV)) and

CC RGD; (2) preparing (A) in which FP contains an HBV core antigen (HBcAg),

CC (P) and RGD; (3) FP; (4) DNA encoding FP; and (5) expression vector

CC containing the DNA of (d). The products of the invention are used in gene

CC therapy of cells and tissues, in vivo or ex vivo. This sequence

CC represents a fusion protein which is described in the method of the

CC invention.

XX Sequence 215 AA;

Query Match 100.0%; Score 61; DB 21; Length 215;

Best Local Similarity 100.0%; Pred. No. 0.00086; Mismatches 0; Indels 0; Gaps 0;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSISFSTRIGDP 12

DB 2 PLSISFSTRIGDP 13

RESULT 12

AAR27471 AAR27471 standard; Protein; 281 AA.

XX AAR27471;

XX 24-FEB-1993 (first entry)

DE spsAg protein.

XX Hepatitis B virus; HBV; M protein; small pre-S antigen; spsAg;

KM vaccinia virus; H6; early/late; promoter; NYVAC; recombinant; HBV L;

KM large pre-S antigen; lpsAg; fusion protein; pre-S region; S12/core;

KM S1; S2; Copenhagen vaccine strain; vaccinia virus; virulence factor;

XX deletion loci; recipient loci.

OS Synthetic.

PN MO9215672-A.

XX 17-SEP-1992.

XX 09-MAR-1992; 92WO-US01906.

XX 07-MAR-1991; 91US-0666056.

XX 11-JUN-1991; 91US-0713967.

XX 06-MAR-1992; 92US-0847951.

XX (VIRIO-) VIROGENETICS CORP.

XX Cok WI, De Taiene C, Francis J, Gettig RR, Johnson GP;

PI Lindach KJ, Norton CK, Paolucci E, Perkins ME, Plincus SR;

PI Riviere M, Tartaglia J, Taylor J;

XX WPI; 1992-331718/40.

DR N-PSDB; AAQ29103.

XX Vaccine comprising recombinant, attenuated pox-virus - use for

PT vaccinating against viral infections such as rabies, hepatitis B,

PT HIV, HSV, EBV, CMV, mumps etc.

XX Disclosure; Fig 9; 456pp; English.

XX The sequence given is encoded by an expression cassette which

CC consists of the hepatitis B virus (HBV) M protein (small pre-S

CC antigen, spsAg) gene precisely linked to a modified synthetic

CC vaccinia virus H6 early/late promoter. This DNA sequence was used in

CC the construction of a NYVAC recombinant expressing the HBV gene.

CC Other HBV genes were also used in the construction. These were HBV

CC L (large pre-S antigen, lpsAg) and a fusion protein composed of the

CC entire pre-S region (S12/core, S1 + S2). Each of these gene sequences

CC were inserted individually into three different sites of NYVAC

CC separated by from each other by large regions of vaccinia DNA

CC containing essential genes. NYVAC is a Copenhagen vaccine strain of

CC vaccinia virus which has been modified by deletion of six non-essential

CC regions of the genome encoding known or potential virulence factors

CC of foreign genes. The spacing of the three inserted sequences ensured

CC that any recombination that did occur would lead to disruption of the

CC vaccinia genome and would cause unviable vaccinia virus. See also

CC AAQ35501-864.

XX Sequence 281 AA;
 Query Match 100.0%; Score 61; DB 13; Length 281;
 Best Local Similarity 100.0%; Pred. No. 0.0011;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 PLSSIFSRIGDP 12
 41 PLSSIFSRIGDP 52

RESULT 13

AAV32835
 ID AAV32835 standard; Protein; 281 AA.

AC AAV32835;

DT 29-OCT-1999 (first entry)

DE HBsAg PreS2-S region protein sequence.

XX HBsAg; PreS2-S; recombinant antigen library; disease-related antigen;
 XX multivalent antigenic polypeptide production; infection; allergen;
 XX asthma; autoimmune disease; rheumatoid arthritis; diabetes; therapy;
 XX multiple sclerosis; inflammatory condition; cancer; contraception;
 XX immune response; hepatitis B surface antigen.

OS Hepatitis B virus.

PN WO9941383-A1.

PD 19-AUG-1999.

PR 10-FEB-1999; 99WO-US02944.

PR 23-OCT-1998; 98US-0105509.

PR 11-FEB-1998; 98US-0021769.

PR 11-FEB-1998; 98US-0074294.

PA (MAXY-) MAXYGEN INC.

PI Base SH, Howard R, Punnonen J, Stemmer WPC, Whalen RG;

DR WPI, 1999-518452/43.

PT Recombinant multivalent antigenic polypeptide produced by

PS against e.g. infections and cancer

Example 14; Fig 17; 153pp; English.

XX This sequence is the hepatitis B virus (HBV) surface antigen (HBsAg)

XX PreS2-S region. This sequence was used to create a recombinant antigen

XX library. The library comprises recombinant nucleic acids encoding

XX antigenic polypeptides and is produced by recombinant nucleic acids encoding

XX forms of nucleic acid, differing by at least two nucleotides, encoding a

XX disease-related antigenic polypeptide. The library can be used to produce

XX a recombinant multivalent antigenic polypeptides (AD) from different

XX polypeptides. The multivalent antigenic polypeptides are used in vaccines

XX to induce a protective or therapeutic response to a wide variety of

XX infectious agents (bacteria, viruses, parasites, including Plasmodium

XX falciparum); allergens; asthma; autoimmune disease (e.g. rheumatoid

XX arthritis, diabetes, multiple sclerosis); other inflammatory conditions

XX for cancer, also, where directed against sperm antigens, they can be used

XX optimised immune response against a wide variety of antigens, they can be used

XX particularly a broad spectrum response to many different strains of a

XX pathogen, including strains that are likely to appear in the future.

XX Query Match 100.0%; Score 61; DB 20; Length 281;
 XX Best Local Similarity 100.0%; Pred. No. 0.0011;
 XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX 1 PLSSIFSRIGDP 12
 XX 41 PLSSIFSRIGDP 52

RESULT 14

AAAR27472
 ID AAAR27472 standard; Protein; 389 AA.

AC AAAR27472;

DT 24-FEB-1993 (first entry)

DE lpsAg protein.

XX Hepatitis B virus; HBV; M protein; small pre-S antigen; spsAg;
 XX vaccinia virus; hemorrhagic region; u/ promoter; NYVAC; recombinant;
 XX HBV L; large pre-S antigen; lpsAg; fusion protein; pre-S region;
 XX S12/core; S1; S2; Copenhagen vaccine strain; vaccinia virus;
 XX virulence factor; deletion loci; recipient loci.

OS Synthetic.

PN Key Location/Qualifiers

FT Region 1..108

FT Region /label S1

FT Region /label S2

FT misc_RNA 164..389

FT /label= S

PN WO9215672-A.

PD 17-SEP-1992.

PR 09-MAR-1992; 92WO-US01906.

PR 07-MAR-1991; 91US-066056.

PR 11-JUN-1991; 91US-0713967.

PR 06-MAR-1992; 92US-0847951.

PA (VIRCO-) VIRGENETICS CORP.

PI Cox WI, De Taisne C, Francis J, Gettig RR, Johnson GP;

DR Limbach KJ, Norton EK, Paolletti E, Perkins ME, Plinius SE;

DR Riviere M, Tartaglia J, Taylor J;

DR WPI, 1992-331718/40.

PT N-PSDB; AAQ29104.

PT Vaccine comprises recombinant, attenuated pox-virus - use for

PT HIV, HSV, BEV, CMV, mumps etc.

PS Disclosure; Fig 11; 456pp; English.

XX The sequence given is encoded by an expression vector which comprises

XX the hepatitis B virus (HBV) L protein (large pre-S antigen, lpsAg)

XX gene linked to the complex hemorrhagic region (u) promoter. This DNA

XX sequence was used in the construction of a NYVAC recombinant

XX expressing the HBV gene. Other HBV genes were also used in the

XX construction. These were HBV M protein (small pre-S antigen, spsAg)

XX and a fusion protein composed of the entire pre-S region (S12/core,

XX S1 + S2). Each of these gene sequences were inserted individually

XX into three different sites of NYVAC separated by from each other by

XX large regions of vaccinia DNA containing essential genes. NYVAC is a

XX Copenhagen vaccine strain of vaccinia virus which has been modified by

XX deletion of six non-essential regions of the genome encoding known or

CC potential virulence factors. The deletion loci were engineered as
 CC recipient loci for the insertion of foreign genes. The spacing of the
 CC three inserted sequences ensured that any recombination that did occur
 CC would lead to disruption of the vaccinia genome and would cause
 CC unviable vaccinia virus. See also AAQ35501-864.

XX Sequence 389 AA;

Query Match 100.0%; Score 61; DB 13; Length 389;

Best Local Similarity 100.0%; Pred. No. 0.0016;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PLSSTFSRIGDP 12

DB 149 PLSSTFSRIGDP 160

RESULT 15

AA27474 standard; Protein; 389 AA.

AA27474;

24-EBB-1993 (first entry)

13L promoter/S12/core gene.

Hepatitis B virus; HBV; M protein; small pre-S antigen; spsAg;

vaccinia virus; Amascta moorei entomopoxvirus; AMEPV; 42 kD; promoter;

NYVAC; recombinant; HBV L; large pre-S antigen; lpsAg; fusion protein;

pre-S region; S12/core; S1; S2; Copenhagen vaccine strain;

vaccinia virus; virulence factor; deletion loci; recipient loci.

Synthetic.

Key Location/Qualifiers

Region 1..108

Region 109..163

Region 164..389

Region /label= S

MO9215672-A.

17-SEP-1992.

09-MAR-1992; 92WO-US01906.

07-MAR-1991; 91US-0666056.

11-JUN-1991; 91US-0713967.

06-MAR-1992; 92US-0847951.

(VIRO-) VIROGENETICS CORP.

COX WT, De Taisne C, Francis J, Gettig RR, Johnson GP;

Limbach KJ, Norton BK, Paolucci E, Perkins ME, Pincus SE;

Riviere M, Tartaglia J, Taylor J;

WPI; 1992-331718/40.

N-PSDB; AAQ29106.

Vaccine comprises recombinant, attenuated pox-virus - use for

vaccinating against viral infections such as rabies, hepatitis B,

HIV, HSV, EBV, CMV, mumps etc.

Disclosure; Fig 15; 456pp; English.

CC Other HBV genes were also used in the construction. These were HBV M
 CC protein (small pre-S antigen, spsAg) and a fusion protein composed of
 CC the entire pre-S region (S12/core, S1 + S2). Each of these gene
 CC sequences were inserted individually into three different sites of
 CC NYVAC separated by from each other by large regions of vaccinia DNA
 CC containing essential genes. NYVAC is a Copenhagen vaccine strain of
 CC vaccinia virus which has been modified by deletion of six non-essential
 CC regions of the genome encoding known or potential virulence factors.
 CC The deletion loci were engineered as recipient loci for the insertion
 CC of foreign genes. The spacing of the three inserted sequences ensured
 CC that any recombination that did occur would lead to disruption of the
 CC vaccinia genome and would cause unviable vaccinia virus. See also
 CC AAQ35501-864.

XX Sequence 389 AA;

Query Match 100.0%; Score 61; DB 13; Length 389;

Best Local Similarity 100.0%; Pred. No. 0.0016;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PLSSTFSRIGDP 12

DB 149 PLSSTFSRIGDP 160

Search completed: February 3, 2003, 09:23:52
 Job time : 2.92335 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd

OM protein - protein search, using sw model

Run on: February 3, 2003, 09:23:57 ; Search time 8.57143 Seconds

(without alignments)
902.626 Million cell updates/sec

Title: US-09-890-752A-20
Perfect score: 61

Scoring table: BIOSUM62

Searched: 4569144 beqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```

Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

```

```

27: pending Patents AAA1n1.*
1: /cgn2_6/ptodataa/1/paa/PCURS COMB. pep.*
2: /cgn2_6/ptodataa/1/paa/US06 COMB. pep.*
3: /cgn2_6/ptodataa/1/paa/US07 COMB. pep.*
4: /cgn2_6/ptodataa/1/paa/US08 COMB. pep.*
5: /cgn2_6/ptodataa/1/paa/US08 COMB. pep.*
6: /cgn2_6/ptodataa/1/paa/US08 COMB. pep.*
7: /cgn2_6/ptodataa/1/paa/US08 COMB. pep.*
8: /cgn2_6/ptodataa/1/paa/US08 COMB. pep.*
9: /cgn2_6/ptodataa/1/paa/US08 COMB. pep.*
10: /cgn2_6/ptodataa/1/paa/US08 COMB. pep.*
11: /cgn2_6/ptodataa/1/paa/US08 COMB. pep.*
12: /cgn2_6/ptodataa/1/paa/US08 COMB. pep.*
13: /cgn2_6/ptodataa/1/paa/US09 COMB. pep.*
14: /cgn2_6/ptodataa/1/paa/US09 COMB. pep.*
15: /cgn2_6/ptodataa/1/paa/US09 COMB. pep.*
16: /cgn2_6/ptodataa/1/paa/US09 COMB. pep.*
17: /cgn2_6/ptodataa/1/paa/US09 COMB. pep.*
18: /cgn2_6/ptodataa/1/paa/US09 COMB. pep.*
19: /cgn2_6/ptodataa/1/paa/US09 COMB. pep.*
20: /cgn2_6/ptodataa/1/paa/US09 COMB. pep.*
21: /cgn2_6/ptodataa/1/paa/US09 COMB. pep.*
22: /cgn2_6/ptodataa/1/paa/US09 COMB. pep.*
23: /cgn2_6/ptodataa/1/paa/US09 COMB. pep.*
24: /cgn2_6/ptodataa/1/paa/US10 COMB. pep.*
25: /cgn2_6/ptodataa/1/paa/US10 COMB. pep.*
26: /cgn2_6/ptodataa/1/paa/US10 COMB. pep.*
27: /cgn2_6/ptodataa/1/paa/US10 COMB. pep.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	61	100.0	12	22	US-09-830-981-2	Sequence 2, Appl1
2	61	100.0	12	22	US-09-830-981-4	Sequence 2, Appl1
3	61	100.0	12	22	US-09-830-981-12	Sequence 14, Appl1
4	61	100.0	12	22	US-09-830-981-14	Sequence 14, Appl1
5	61	100.0	12	22	US-09-830-981-16	Sequence 16, Appl1
6	61	100.0	12	22	US-09-830-981-18	Sequence 18, Appl1

7	61	100.0	12	22	US-09-830-751-20	Sequence 20, App
8	61	100.0	12	22	US-09-890-582A-20	Sequence 20, App
9	61	100.0	12	24	US-10-077-555-9	Sequence 9, App1
10	61	100.0	12	25	US-10-144-549-12	Sequence 12, App
11	61	100.0	12	26	US-10-226-956-284	Sequence 284, App
12	61	100.0	102	14	US-09-021-769-45	Sequence 45, App
13	61	100.0	102	16	US-09-248-716-6	Sequence 6, App1
14	61	100.0	102	21	US-09-724-890-45	Sequence 45, App1
15	61	100.0	164	1	PCT-US99-31020-41	Sequence 41, App1
16	61	100.0	164	18	US-09-471-573A-41	Sequence 41, App1
17	61	100.0	215	22	US-09-890-752A-2	Sequence 41, App1
18	61	100.0	281	3	US-07-847-951A-214	Sequence 214, App
19	61	100.0	281	4	US-08-036-218-112	Sequence 212, App
20	61	100.0	281	14	US-09-021-769-6	Sequence 6, App1
21	61	100.0	281	16	US-09-247-890-12	Sequence 12, App1
22	61	100.0	281	21	US-09-724-648-12	Sequence 12, App1
23	61	100.0	281	21	US-09-724-648-12	Sequence 12, App1
24	61	100.0	281	21	US-09-724-648-12	Sequence 12, App1
25	61	100.0	389	3	US-07-847-951A-215	Sequence 215, App
26	61	100.0	389	4	US-08-036-218-114	Sequence 214, App
27	61	100.0	389	4	US-08-036-218-117	Sequence 217, App
28	61	100.0	389	7	US-08-360-107-116	Sequence 116, App
29	61	100.0	389	8	US-08-470-896-106	Sequence 106, App
30	61	100.0	389	8	US-08-471-913-106	Sequence 106, App
31	61	100.0	389	8	US-08-475-668-106	Sequence 106, App
32	61	100.0	389	8	US-08-484-223A-106	Sequence 106, App
33	61	100.0	389	8	US-08-485-546-106	Sequence 106, App
34	61	100.0	389	8	US-08-485-546A-106	Sequence 106, App
35	61	100.0	389	8	US-08-485-551-106	Sequence 106, App
36	61	100.0	389	8	US-08-487-266-106	Sequence 106, App
37	61	100.0	389	8	US-08-487-266A-106	Sequence 106, App
38	61	100.0	389	8	US-08-487-266-106	Sequence 106, App
39	61	100.0	389	8	US-08-487-266A-106	Sequence 106, App
40	61	100.0	389	8	US-08-487-266-106	Sequence 106, App
41	61	100.0	389	8	US-08-487-266A-106	Sequence 106, App
42	61	100.0	389	13	US-08-919-600-106	Sequence 106, App
43	61	100.0	389	19	US-09-502-445-106	Sequence 106, App
44	61	100.0	389	26	US-10-267-689-106	Sequence 106, App
45	61	100.0	389	26	US-10-267-746-106	Sequence 106, App

ALIGNMENTS

```

1 RESULT 1
2 US-09-830-981-2
3 Sequence 2, Application US/09830981
4 GENERAL INFORMATION:
5 APPLICANT: Eberhard Hildt
6 APPLICANT: Stephanie Schmidt
7 TITLE OF INVENTION: Polypeptide Mediating Cell Permeability
8 FILE REFERENCE: 033392-001
9 CURRENT APPLICATION NUMBER: US/09/830,981
10 CURRENT FILING DATE: 2002-04-15
11 PRIOR APPLICATION NUMBER: PCT/DE99/03506
12 PRIOR FILING DATE: 1999-11-03
13 PRIOR APPLICATION NUMBER: DE 199 50 718.6
14 PRIOR FILING DATE: 1998-11-03
15 NUMBER OF SEQ ID NOS: 20
16 SOFTWARE: PatentIn Ver. 2.1
17 SEQ ID NO 2
18 LENGTH: 12
19 TYPE: PRT
20 ORGANISM: Artificial Sequence
21 FEATURE:
22 OTHER INFORMATION: Description of Artificial Sequence:
23 OTHER INFORMATION: Cell permeability mediating polypeptide
24 US-09-830-981-2

```

QY 1 PLSISFRIIDP 12
| | | | | | | | | |
Db 1 PLSISFRIIDP 12

RESULT 2
US-09-830-981-14
Sequence 4, Application US/09830981
GENERAL INFORMATION:
APPLICANT: Eberhard Hildt
APPLICANT: Stephanie Schmidt
TITLE OF INVENTION: Polypeptide Mediating Cell Permeability
FILE REFERENCE: 033392-001
CURRENT FILING DATE: 2002-04-15
PRIOR APPLICATION NUMBER: PCT/DE99/03506
PRIOR FILING DATE: 1999-11-03
PRIOR APPLICATION NUMBER: DE 198 50 718.6
PRIOR FILING DATE: 1998-11-03
NUMBER OF SEQ ID NOS: 20
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 4
LENGTH: 12
TYPE: PRT
ORGANISM: Hepadnavirus
US-09-830-981-14

Query Match 100.0%; Score 61; DB 22; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00042;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFRIIDP 12
| | | | | | | | | |
Db 1 PLSISFRIIDP 12

RESULT 3
US-09-830-981-12
Sequence 12, Application US/09830981
GENERAL INFORMATION:
APPLICANT: Eberhard Hildt
APPLICANT: Stephanie Schmidt
TITLE OF INVENTION: Polypeptide Mediating Cell Permeability
FILE REFERENCE: 033392-001
CURRENT FILING DATE: 2002-04-15
PRIOR APPLICATION NUMBER: PCT/DE99/03506
PRIOR FILING DATE: 1999-11-03
PRIOR APPLICATION NUMBER: DE 198 50 718.6
PRIOR FILING DATE: 1998-11-03
NUMBER OF SEQ ID NOS: 20
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 12
LENGTH: 12
TYPE: PRT
ORGANISM: Hepadnavirus
US-09-830-981-12

Query Match 100.0%; Score 61; DB 22; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00042;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFRIIDP 12
| | | | | | | | | |
Db 1 PLSISFRIIDP 12

RESULT 4
US-09-830-981-14
Sequence 14, Application US/09830981
GENERAL INFORMATION:
APPLICANT: Eberhard Hildt
APPLICANT: Stephanie Schmidt

TITLE OF INVENTION: Polypeptide Mediating Cell Permeability
FILE REFERENCE: 033392-001
CURRENT FILING DATE: 2002-04-15
PRIOR APPLICATION NUMBER: PCT/DE99/03506
PRIOR FILING DATE: 1999-11-03
PRIOR APPLICATION NUMBER: DE 198 50 718.6
PRIOR FILING DATE: 1998-11-03
NUMBER OF SEQ ID NOS: 20
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 14
LENGTH: 12
TYPE: PRT
ORGANISM: Hepadnavirus
US-09-830-981-14

Query Match 100.0%; Score 61; DB 22; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00042;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFRIIDP 12
| | | | | | | | | |
Db 1 PLSISFRIIDP 12

RESULT 5
US-09-830-981-16
Sequence 16, Application US/09830981
GENERAL INFORMATION:
APPLICANT: Eberhard Hildt
APPLICANT: Stephanie Schmidt
TITLE OF INVENTION: Polypeptide Mediating Cell Permeability
FILE REFERENCE: 033392-001
CURRENT FILING DATE: 2002-04-15
PRIOR APPLICATION NUMBER: PCT/DE99/03506
PRIOR FILING DATE: 1999-11-03
PRIOR APPLICATION NUMBER: DE 198 50 718.6
PRIOR FILING DATE: 1998-11-03
NUMBER OF SEQ ID NOS: 20
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 16
LENGTH: 12
TYPE: PRT
ORGANISM: Hepadnavirus
US-09-830-981-16

Query Match 100.0%; Score 61; DB 22; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00042;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFRIIDP 12
| | | | | | | | | |
Db 1 PLSISFRIIDP 12

RESULT 6
US-09-830-981-18
Sequence 18, Application US/09830981
GENERAL INFORMATION:
APPLICANT: Eberhard Hildt
APPLICANT: Stephanie Schmidt
TITLE OF INVENTION: Polypeptide Mediating Cell Permeability
FILE REFERENCE: 033392-001
CURRENT FILING DATE: 2002-04-15
PRIOR APPLICATION NUMBER: PCT/DE99/03506
PRIOR FILING DATE: 1999-11-03
PRIOR APPLICATION NUMBER: DE 198 50 718.6
PRIOR FILING DATE: 1998-11-03
NUMBER OF SEQ ID NOS: 20
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 18

LENGTH: 12
TYPE: PRT
ORGANISM: Hepadnavirus
US-09-830-981-18

Query Match 100.0%; Score 61; DB 22; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00042;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFRIQDP 12
DB 1 PLSISFRIQDP 12

RESULT 7
US-09-830-981-20

Sequence 20, Application US/09830981
GENERAL INFORMATION:
APPLICANT: Eberhard Hildt
APPLICANT: Stephanle Schmidt
TITLE OF INVENTION: Polypeptide Mediating Cell Permeability
FILE REFERENCE: 033392-001
CURRENT APPLICATION NUMBER: US/09/830,981
CURRENT FILING DATE: 2002-04-15
PRIOR APPLICATION NUMBER: PCT/DE99/03506
PRIOR FILING DATE: 1999-11-03
PRIOR APPLICATION NUMBER: DE 198 50 718.6
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 20
LENGTH: 12
TYPE: PRT
ORGANISM: Hepadnavirus
US-09-830-981-20

Query Match 100.0%; Score 61; DB 22; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00042;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFRIQDP 12
DB 1 PLSISFRIQDP 12

RESULT 8
US-09-890-752A-20

Sequence 20, Application US/09890752A
GENERAL INFORMATION:
APPLICANT: Hildt, Eberhard
APPLICANT: Hofschneider, Peter
TITLE OF INVENTION: Particles for Gene Therapy
FILE REFERENCE: 107070-120 (VOS-013)
CURRENT APPLICATION NUMBER: US/09/890,752A
CURRENT FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: PCT/DE00/00363
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: DE 199 04 800.2
NUMBER OF SEQ ID NOS: 21
SOFTWARE: PatentIn version 3.1
SEQ ID NO 20
LENGTH: 12
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: cell permeability-mediating peptide
US-09-890-752A-20

Query Match 100.0%; Score 61; DB 22; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00042;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFRIQDP 12
DB 1 PLSISFRIQDP 12

RESULT 9
US-10-077-555-9

Sequence 9, Application US/10077555
GENERAL INFORMATION:
APPLICANT: Wang, Rong-fu
TITLE OF INVENTION: Use of Cell-Penetrating Peptides to Generate Antitumor Immunity
FILE REFERENCE: P02373US/10200806
CURRENT APPLICATION NUMBER: US/10/077,555
CURRENT FILING DATE: 2002-02-15
PRIOR APPLICATION NUMBER: US 60/266,687
PRIOR FILING DATE: 2001-02-15
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn version 3.1
SEQ ID NO 9
LENGTH: 12
TYPE: PRT
ORGANISM: Hepatitis B virus
US-10-077-555-9

Query Match 100.0%; Score 61; DB 24; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00042;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFRIQDP 12
DB 1 PLSISFRIQDP 12

RESULT 10
US-10-144-549-12

Sequence 12, Application US/10144549
GENERAL INFORMATION:
APPLICANT: Geneshuttle Biopharm, Inc.
APPLICANT: Hwu, Paul L.
TITLE OF INVENTION: A NEW FUSION PROTEIN FOR USE AS VECTOR
FILE REFERENCE: MEHB 02-340
CURRENT APPLICATION NUMBER: US/10/144,549
CURRENT FILING DATE: 2002-05-13
NUMBER OF SEQ ID NOS: 31
SOFTWARE: PatentIn version 3.1
SEQ ID NO 12
LENGTH: 12
TYPE: PRT
ORGANISM: Hepatitis B virus
FEATURE:
NAME/KEY: MISC FEATURE
OTHER INFORMATION: The HBV PreS2 antigen consisting of the translocation motif from
US-10-144-549-12

Query Match 100.0%; Score 61; DB 25; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00042;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFRIQDP 12
DB 1 PLSISFRIQDP 12

RESULT 11
US-10-226-956-284

Sequence 284, Application US/10226956
GENERAL INFORMATION:
APPLICANT: Brophy, Colleen
APPLICANT: Komalavilas, Padmini
APPLICANT: Panitch, Alyssa
APPLICANT: Joshi, Lokesh
APPLICANT: Seal, Brandon L.

TITLE OF INVENTION: REAGENTS AND METHODS FOR SMOOTH MUSCLE THERAPIES
FILE REFERENCE: ASU-1061-US
CURRENT APPLICATION NUMBER: US/10/226,956
CURRENT FILING DATE: 2002-08-23
PRIOR APPLICATION NUMBER: 60/314,535
PRIOR FILING DATE: 2001-08-23
NUMBER OF SEQ ID NOS: 320
SOFTWARE: PatentIn version 3.1
SEQ ID NO 284
LENGTH: 12
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Synthetic peptide
US-10-226-956-284

Query Match 100.0%; Score 61; DB 26; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00042;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSIFSRIGDP 12
Db 1 PLSIFSRIGDP 12

RESULT 12
US-09-021-769-45
Sequence 45, Application US/09021769
GENERAL INFORMATION:
APPLICANT: Punnonen, Juha
APPLICANT: Stemmer, Willem P.C.
APPLICANT: Whalen, Robert Gerald
APPLICANT: Howard, Russell
TITLE OF INVENTION: Methods and Compositions for Vaccine
TITLE OF INVENTION: Engineering
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSER: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/021,769
FILING DATE: 11-FEB-1998
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Timothy L.
REGISTRATION NUMBER: 35,367
REFERENCE/DOCKET NUMBER: 018097-021500US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 102 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..102
OTHER INFORMATION: /notes "truncated Hepatitis B surface
OTHER INFORMATION: antigen (HbsAg) (Pres2 plus S regions)
OTHER INFORMATION: polypeptide"
US-09-021-769-45

Query Match 100.0%; Score 61; DB 14; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSIFSRIGDP 12
Db 41 PLSIFSRIGDP 52

RESULT 13
US-09-248-716-6
Sequence 6, Application US/09248716
GENERAL INFORMATION:
APPLICANT: Punnonen, Juha
APPLICANT: Stemmer, Willem P.C.
APPLICANT: Whalen, Robert Gerald
APPLICANT: Howard, Russell
APPLICANT: Maxygen, Inc.
TITLE OF INVENTION: Optimization of Immunomodulatory Properties of Genetic
TITLE OF INVENTION: Vaccines
FILE REFERENCE: 018097-030300US
CURRENT APPLICATION NUMBER: US/09/248,716
CURRENT FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: US 60/074,294
PRIOR FILING DATE: 1998-02-11
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 6
LENGTH: 102
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: truncated
OTHER INFORMATION: hepatitis B surface antigen (HbsAg) (Pres2 plus S
OTHER INFORMATION: regions)
US-09-248-716-6

Query Match 100.0%; Score 61; DB 16; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSIFSRIGDP 12
Db 41 PLSIFSRIGDP 52

RESULT 14
US-09-724-890-45
Sequence 45, Application US/09724890
GENERAL INFORMATION:
APPLICANT: Punnonen, Juha
APPLICANT: Stemmer, Willem P.C.
APPLICANT: Whalen, Robert Gerald
APPLICANT: Howard, Russell
TITLE OF INVENTION: Methods and Compositions for Vaccine
TITLE OF INVENTION: Engineering
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSER: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/724,890
FILING DATE: 28-NOV-2000

```

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/021,769
FILING DATE: 11-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Timothy L.
REGISTRATION NUMBER: 35,367
REFERENCE/DOCKET NUMBER: 018097-021500US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 102 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..102
OTHER INFORMATION: /note="truncated Hepatitis B
antigen (HBeAg) (Pres2 plus S
regions)
polypeptide"
SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-09-724-890-45

```

```

Query Match          100.0%; Score 61; DB 1; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 PLSSTPSRIGDP 12
    |||||
Db 41 PLSSTPSRIGDP 52

```

```

RESULT 15
PCT-US99-31020-41
; Sequence 41, Application PC/TUS9931020
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Inst. for Plant Res. at Cornell
; TITLE OF INVENTION: Expression of Immunogenic Hepatitis B Surface Antigens
; TITLE OF INVENTION: in Transgenic Plants
; FILE REFERENCE: 4868/85428
; CURRENT APPLICATION NUMBER: PCT/US99/31020
; CURRENT FILING DATE: 2000-04-10
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Plant-optimized
; OTHER INFORMATION: pre S1/S2.
PCT-US99-31020-41

```

```

Query Match          100.0%; Score 61; DB 1; Length 164;
Best Local Similarity 100.0%; Pred. No. 0.0074;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PLSSTPSRIGDP 12
    |||||
Db 149 PLSSTPSRIGDP 160

```

```

Search completed: February 3, 2003, 09:34:08
Job time: 8.57143 secs

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 09:21:12 ; Search time 1.60976 Seconds
(without alignments)
1535.988 Million cell updates/sec

Title: US-09-890-752a-20
Perfect score: 61
Sequence: 1 PUSIFSRIQDP 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21: *
1: sp_archaea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mhc: *
8: sp_organelle: *
9: sp_phage: *
10: sp_plant: *
11: sp_protent: *
12: sp_virus: *
13: sp_vertebrate: *
14: sp_unclassified: *
15: sp_virus: *
16: sp_bacterioph: *
17: sp_archaeop: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61	100.0	57	12 091IC1	091IC1 hepatitis b
2	61	100.0	163	12 09WP19	09WP19 hepatitis b
3	61	100.0	281	12 041750	041750 hepatitis b
4	61	100.0	281	12 08QXP8	08QXP8 hepatitis b
5	61	100.0	378	12 067944	067944 hepatitis b
6	61	100.0	378	12 067954	067954 hepatitis b
7	61	100.0	383	12 096837	096837 hepatitis b
8	61	100.0	383	12 096840	096840 hepatitis b
9	61	100.0	389	12 090772	090772 hepatitis b
10	61	100.0	389	12 09WMX3	09WMX3 hepatitis b
11	61	100.0	389	12 091C52	091C52 hepatitis b
12	61	100.0	389	12 091C55	091C55 hepatitis b
13	61	100.0	389	12 067875	067875 hepatitis b
14	61	100.0	389	12 08QXP9	08QXP9 hepatitis b
15	59	96.7	49	12 091IC3	091IC3 hepatitis b
16	59	96.7	55	12 091XFA	091XFA hepatitis b

17	59	96.7	55	12 091XF2	091XF2 hepatitis b
18	59	96.7	55	12 091XF1	091XF1 hepatitis b
19	59	96.7	55	12 091XE9	091XE9 hepatitis b
20	59	96.7	55	12 091XE8	091XE8 hepatitis b
21	59	96.7	55	12 091XE6	091XE6 hepatitis b
22	59	96.7	55	12 091XE2	091XE2 hepatitis b
23	59	96.7	55	12 091XE0	091XE0 hepatitis b
24	59	96.7	55	12 091XD9	091XD9 hepatitis b
25	59	96.7	55	12 091XD7	091XD7 hepatitis b
26	59	96.7	55	12 091XD6	091XD6 hepatitis b
27	59	96.7	55	12 091XD4	091XD4 hepatitis b
28	59	96.7	55	12 091XD3	091XD3 hepatitis b
29	59	96.7	55	12 091XD1	091XD1 hepatitis b
30	59	96.7	55	12 091XD0	091XD0 hepatitis b
31	59	96.7	55	12 091XC8	091XC8 hepatitis b
32	59	96.7	55	12 091XC7	091XC7 hepatitis b
33	59	96.7	55	12 091XC5	091XC5 hepatitis b
34	59	96.7	55	12 091XC4	091XC4 hepatitis b
35	59	96.7	55	12 091XC2	091XC2 hepatitis b
36	59	96.7	55	12 091XB5	091XB5 hepatitis b
37	59	96.7	55	12 091XB3	091XB3 hepatitis b
38	59	96.7	55	12 091XB2	091XB2 hepatitis b
39	59	96.7	55	12 091XB0	091XB0 hepatitis b
40	59	96.7	55	12 091XA9	091XA9 hepatitis b
41	59	96.7	55	12 091XA7	091XA7 hepatitis b
42	59	96.7	55	12 091XA6	091XA6 hepatitis b
43	59	96.7	55	12 091XA4	091XA4 hepatitis b
44	59	96.7	55	12 091XA3	091XA3 hepatitis b
45	59	96.7	55	12 091XA1	091XA1 hepatitis b

ALIGNMENTS

RESULT 1

ID 091IC1 PRELIMINARY; PRT; 57 AA.
AC 091IC1;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Pre-S2 protein (Fragment).
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N314-1;
RA Huangfu J., Dong J., Deng H.;
RT "Preliminary Study on Pres2 Region of Hepatitis B Virus."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF393223; AAK84362.1; -
DR InterPro: IPR000349; Hepvir_surfAg.
DR Pfam: PF00695; VMSA; 1.
FT NON_TER
SQ SEQUENCE 57 AA; 6124 MW; 8D756AF55FF849CF CRC64;

Query Match 100.0%; Score 61; DB 12; Length 57;
Best Local Similarity 100.0%; Pred. No. 0.00019;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PUSIFSRIQDP 12
|||
Db 41 PUSIFSRIQDP 52

RESULT 2

ID 09WP19 PRELIMINARY; PRT; 163 AA.
AC 09WP19;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

```

DE Large S protein (Fragment).
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OK NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MF;
RA Kidd-Ljunggren K.;
RT "Nosocomial transmission of Hepatitis B Virus Through Multiple Dose
RL Vials";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF136461; AAD4711.1; -
DR InterPro: IPR000349; Hepvir_surflag.
DR Pfam; PF00695; VMSA; 1.
FT NON_TER 163
SQ SEQUENCE 163 AA; 17331 MW; 38D20BE46F0BD24 CRC64;

Query Match 100.0%; Score 61; DB 12; Length 163;
Best Local Similarity 100.0%; Pred. No. 0.00061;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PLSSTFSRIGDP 12
DB 149 PLSSTFSRIGDP 160

RESULT 3
ID 041750 PRELIMINARY; PRT; 281 AA.
AC 041750;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Middle S protein.
GN PRES2/S.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OK NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=808;
RX MEDLINE=97368435; PubMed=9225049;
RA Bowyer S., van Staden L., Kew M.C., Sim J.G.;
RT "A unique segment of the hepatitis B virus group A genotype identified
RT in isolates from South Africa.";
RL J. Gen. Virol. 78:1719-1729(1997).
DR EMBL; U87737; AAC58018.1; -
DR InterPro: IPR000349; Hepvir_surflag.
DR Pfam; PF00695; VMSA; 1.
SQ SEQUENCE 281 AA; 31153 MW; 85D90E7B5F2C9289 CRC64;

Query Match 100.0%; Score 61; DB 12; Length 281;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PLSSTFSRIGDP 12
DB 41 PLSSTFSRIGDP 52

RESULT 4
ID 080XFB PRELIMINARY; PRT; 281 AA.
AC 080XFB;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Middle surface antigen.
GN PRES2.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OK NCBI_TaxID=10407;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RA Kay A.C.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Jeanet D., Chemin I., Mandrand B., Zoulim F., Trepo C., Kay A.;
RT "Characterization of two Hepatitis B virus populations isolated from a
RT Hepatitis B surface antigen-negative patient.";
RL Hepatology 0:0-0(0).
DR EMBL; AJ344117; CAC87014.1; -
SQ SEQUENCE 281 AA; 31266 MW; 37234BBD949B3B03 CRC64;

Query Match 100.0%; Score 61; DB 12; Length 281;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PLSSTFSRIGDP 12
DB 41 PLSSTFSRIGDP 52

RESULT 5
ID 067944 PRELIMINARY; PRT; 378 AA.
AC 067944;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Surface proteins.
GN PRES1/PRES2.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OK NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.
RA Lai M.E., Mazzoleni A.P., Porru A., Balestrieri A.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; X85254; CA55910.1; -
DR InterPro: IPR000349; Hepvir_surflag.
DR Pfam; PF00695; VMSA; 1.
SQ SEQUENCE 378 AA; 41257 MW; D4370B06FPB625D5 CRC64;

Query Match 100.0%; Score 61; DB 12; Length 378;
Best Local Similarity 100.0%; Pred. No. 0.0015;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PLSSTFSRIGDP 12
DB 149 PLSSTFSRIGDP 160

RESULT 6
ID 067954 PRELIMINARY; PRT; 378 AA.
AC 067954;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE HBV surface proteins (Fragment).
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OK NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.
RA Lai M.E., Mazzoleni A.P., Balestrieri A.;
RT "Sequence analysis of HBV genomes isolated from patients with HBsAg
RT chronic liver disease.";
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; X77309; CA54515.1; -
DR InterPro: IPR000349; Hepvir_surflag.
DR Pfam; PF00695; VMSA; 1.
FT NON_TER 378
SQ SEQUENCE 378 AA; 41356 MW; FC18939606E121CD CRC64;

```

Query Match 100.0%; Score 61; DB 12; Length 378;
Best Local Similarity 100.0%; Pred. No. 0.0015;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 PLSSIFSRIGDP 12
Db 149 PLSSIFSRIGDP 160

RESULT 7
Q96837 PRELIMINARY; PRT; 383 AA.
AC Q96837; (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DE Large S protein (Fragment).
GN PRE-S/S.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxId=10407;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRH;
RX MEDLINE=97049068; PubMed=8893798;
RA Moraes M.T., Gomes S.A., Niel C.;
RT "Sequence analysis of pre-S/S gene of hepatitis B virus strains of
RL Arch. Virol. 141:1767-1773(1996)."
DR EMBL; U55224; AAB47470.1; -
DR InterPro; IPR000349; Hepvir_surfAg.
DR Pfam; PF00695; VMSA; 1.
FT NON TER 383
SQ SEQUENCE 383 AA; 41942 MW; 6B101F3B7471B6B1 CRC64;

Query Match 100.0%; Score 61; DB 12; Length 383;
Best Local Similarity 100.0%; Pred. No. 0.0015;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 PLSSIFSRIGDP 12
Db 149 PLSSIFSRIGDP 160

RESULT 8
Q96840 PRELIMINARY; PRT; 383 AA.
AC Q96840; (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DE Large S protein (Fragment).
GN PRE-S/S.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxId=10407;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRJ;
RX MEDLINE=97049068; PubMed=8893798;
RA Moraes M.T., Gomes S.A., Niel C.;
RT "Sequence analysis of pre-S/S gene of hepatitis B virus strains of
RL Arch. Virol. 141:1767-1773(1996)."
DR EMBL; U55227; AAB47473.1; -
DR InterPro; IPR000349; Hepvir_surfAg.
DR Pfam; PF00695; VMSA; 1.
FT NON TER 383
SQ SEQUENCE 383 AA; 42018 MW; 341AB53B7911BBDC CRC64;

Query Match 100.0%; Score 61; DB 12; Length 383;
Best Local Similarity 100.0%; Pred. No. 0.0015;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 PLSSIFSRIGDP 12
Db 149 PLSSIFSRIGDP 160

RESULT 9
Q90772 PRELIMINARY; PRT; 389 AA.
AC Q90772; (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Surface protein.
GN S GEN3.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxId=10407;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=REF2;
RX MEDLINE=98362187; PubMed=9696878;
RA Grethe S., Monazahian M., Boehme I., Thomassen R.;
RT "Characterization of unusual escape variants of hepatitis B virus
RL J. Virol. 72:7692-7696(1998)."
DR EMBL; AJ003116; CA403872.1; -
DR InterPro; IPR000349; Hepvir_surfAg.
DR Pfam; PF00695; VMSA; 1.
SQ SEQUENCE 389 AA; 42682 MW; 575E7223D068CDB CRC64;

Query Match 100.0%; Score 61; DB 12; Length 389;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 PLSSIFSRIGDP 12
Db 149 PLSSIFSRIGDP 160

RESULT 10
Q9MX3 PRELIMINARY; PRT; 389 AA.
AC Q9MX3; (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DE Surface antigen.
GN S.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxId=10407;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=99238050; PubMed=10223539;
RA Petzold D.R., Tautz B., Wolf F., Drecher J.;
RT "Infection chains and evolution rates of Hepatitis B virus in cardiac
RL transplant recipients infected nosocomially.";
J. Med. Virol. 58:1-10(1999)."
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX Petzold D.R.;
RL Submitted (Jan-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ131956; CA10540.1; -
DR InterPro; IPR000349; Hepvir_surfAg.
DR Pfam; PF00695; VMSA; 1.
FT CHAIN 164
SQ SEQUENCE 389 AA; 42708 MW; 6AA0B85DFE5B99C CRC64;

Query Match 100.0%; Score 61; DB 12; Length 389;

Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSIFSRIGDP 12
DB 149 PLSIFSRIGDP 160

RESULT 11

Q91C55 PRELIMINARY; PRT; 389 AA.
AC Q91C55;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Large/middle/small S proteins.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=41;
RA MEDLINE=21361467; PubMed=11468728;
RA Owiredu W.K., Kramvis A., Kew M.C.;
RT "Hepatitis B virus DNA in serum of healthy black African adults
RT positive for hepatitis B surface antibody alone: possible association
RT with recombination between genotypes A and D."
RL J. Med. Virol. 64:441-454(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=41;
RA Owiredu W.K.B.A., Kramvis A., Kew M.C.;
RA Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF297619; AAK97179.1; -
DR InterPro; IPR000349; Hepvir_surfAg.
DR Pfam; PF00695; VMSA; 1
SQ SEQUENCE 389 AA; 42580 MW; F4336E04992CE6E2 CRC64;

Query Match 100.0%; Score 61; DB 12; Length 389;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSIFSRIGDP 12
DB 149 PLSIFSRIGDP 160

RESULT 12
Q91C52 PRELIMINARY; PRT; 389 AA.
AC Q91C52;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Large/middle/small S proteins.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=57;
RX MEDLINE=21361467; PubMed=11468728;
RA Owiredu W.K., Kramvis A., Kew M.C.;
RT "Hepatitis B virus DNA in serum of healthy black African adults
RT positive for hepatitis B surface antibody alone: possible association
RT with recombination between genotypes A and D."
RL J. Med. Virol. 64:441-454(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=57;
RA Owiredu W.K.B.A., Kramvis A., Kew M.C.;
RA Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF297620; AAK97183.1; -

DR InterPro; IPR000349; Hepvir_surfAg.
DR Pfam; PF00695; VMSA; 1
SQ SEQUENCE 389 AA; 42937 MW; 1709DFD7D5E01066 CRC64;

Query Match 100.0%; Score 61; DB 12; Length 389;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSIFSRIGDP 12
DB 149 PLSIFSRIGDP 160

RESULT 13
Q67875 PRELIMINARY; PRT; 389 AA.
AC Q67875;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Pre S1/S ORF.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.
RA Lai M.E., Mazzoleni A.P., Balestrieri A., Mellis A., Porru A.;
RT "Sequence analysis of Hbv genomes isolated from patients with HBeAg
RT negative chronic liver disease."
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; X65258; CAA46353.1; -
DR InterPro; IPR000349; Hepvir_surfAg.
DR Pfam; PF00695; VMSA; 1
SQ SEQUENCE 389 AA; 42661 MW; A39542B416E48F24 CRC64;

Query Match 100.0%; Score 61; DB 12; Length 389;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSIFSRIGDP 12
DB 149 PLSIFSRIGDP 160

RESULT 14
Q8QXP9 PRELIMINARY; PRT; 389 AA.
AC Q8QXP9;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Large surface antigen.
CN PRS1.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.
RA Kay A.C.;
RA Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Jeanet D., Chemin I., Mandrand B., Zoulim F., Trepo C., Kay A.;
RT "Characterization of two Hepatitis B virus populations isolated from a
RT Hepatitis B surface antigen-negative patient."
RL Hepatology 0:0-0(0)
DR EMBL; AJ344117; CAC67013.1; -
SQ SEQUENCE 389 AA; 42764 MW; 9AC5A4D46B73632 CRC64;

Query Match 100.0%; Score 61; DB 12; Length 389;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSSIFSRIGDP 12
 |||||
 DB 149 PLSSIFSRIGDP 160

RESULT 15

091IC3
 ID 091IC3 PRELIMINARY; PRT; 49 AA.
 AC 091IC3;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Pre-S2 protein (Fragment).
 OS Hepatitis B virus.
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
 OK NCBI_TaxID=10407;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=702-4;
 RA Huangfu J., Dong J., Deng H.;
 RT "Preliminary Study on Pres2 Region of Hepatitis B Virus."
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF393221; AAK84360.1; -
 DR InterPro; IPR000349; Hepvir_surfa.
 DR Pfam; PF00695; VMSA; 1.
 FT NON_TER 49
 SQ SEQUENCE 49 AA; 5045 MM; 5CA3FDA356B59C3 CRC64;
 Query Match 96.7%; Score 59; DB 12; Length 49;
 Best Local Similarity 91.7%; Pred. No. 0.00041;
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PLSSIFSRIGDP 12
 |||||
 DB 33 PLSSIFSRIGDP 44

Search completed: February 3, 2003, 09:25:45
 Job time : 2.94309 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 09:21:12 ; Search time 46.5488 Seconds
(without alignments)
1535.988 Million cell updates/sec

Title: US-09-890-752a-1

Sequence: 1 MWRGAGAGAGGFTPPHGG.....SLSPFLPLPIFCLWYI 347

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 segs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP virus:*
16: SP bacteriophage:*
17: SP archaea:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1901	98.5	389	12 Q9WJC4	Q9WJC4 hepatitis b
2	1898	98.4	389	12 Q96843	Q96843 hepatitis b
3	1885	97.7	389	12 Q9PWW6	Q9PWW6 hepatitis b
4	1883	97.6	389	12 Q67871	Q67871 hepatitis b
5	1883	97.6	389	12 Q67914	Q67914 hepatitis b
6	1879	97.4	389	12 Q67920	Q67920 hepatitis b
7	1878	97.4	389	12 Q9QWH6	Q9QWH6 hepatitis b
8	1876	97.3	389	12 Q91F39	Q91F39 hepatitis b
9	1875	97.2	389	12 Q67893	Q67893 hepatitis b
10	1875	97.2	389	12 Q8OXQ3	Q8OXQ3 hepatitis b
11	1875	97.2	389	12 Q8OXQ9	Q8OXQ9 hepatitis b
12	1874	97.1	389	12 Q9QAF2	Q9QAF2 hepatitis b
13	1864	96.6	389	12 Q9QAF9	Q9QAF9 hepatitis b
14	1864	96.6	389	12 Q9WMS1	Q9WMS1 hepatitis b
15	1861	96.5	389	12 Q92921	Q92921 hepatitis b
16	1861	96.5	389	12 Q67879	Q67879 hepatitis b

17	1855	96.2	389	12 Q96839	Q96839 hepatitis b
18	1854	96.1	389	12 Q90772	Q90772 hepatitis b
19	1854	96.1	389	12 Q9WXX3	Q9WXX3 hepatitis b
20	1849	95.9	383	12 Q96838	Q96838 hepatitis b
21	1845	95.6	383	12 Q96837	Q96837 hepatitis b
22	1843	95.5	400	12 Q8VJ11	Q8VJ11 hepatitis b
23	1839	95.3	383	12 Q96840	Q96840 hepatitis b
24	1828	94.8	400	12 Q913A6	Q913A6 hepatitis b
25	1827	94.7	385	12 Q92919	Q92919 hepatitis b
26	1825	94.6	389	12 Q91C52	Q91C52 hepatitis b
27	1824	94.6	383	12 Q96836	Q96836 hepatitis b
28	1816	94.1	389	12 Q9QW10	Q9QW10 hepatitis b
29	1816	94.1	389	12 Q91C55	Q91C55 hepatitis b
30	1816	94.1	389	12 Q67890	Q67890 hepatitis b
31	1814	94.0	383	12 Q96841	Q96841 hepatitis b
32	1809.5	93.8	382	12 Q67867	Q67867 hepatitis b
33	1803	93.5	389	12 Q67956	Q67956 hepatitis b
34	1799	93.3	378	12 Q67954	Q67954 hepatitis b
35	1797	93.2	389	12 Q67875	Q67875 hepatitis b
36	1796	93.1	378	12 Q67944	Q67944 hepatitis b
37	1793	92.9	387	12 Q91A30	Q91A30 hepatitis b
38	1790	92.8	399	12 Q9QAW7	Q9QAW7 hepatitis b
39	1788	92.7	343	12 Q91T59	Q91T59 hepatitis b
40	1786	92.6	344	12 Q91TA6	Q91TA6 hepatitis b
41	1781	92.3	399	12 Q91R25	Q91R25 hepatitis b
42	1780	92.3	343	12 Q91T77	Q91T77 hepatitis b
43	1780	92.3	389	12 Q91R27	Q91R27 hepatitis b
44	1773	91.9	389	12 Q9QAW0	Q9QAW0 hepatitis b
45	1772	91.9	400	12 Q95E78	Q95E78 hepatitis b

ALIGNMENTS

RESULT 1

Q9WJC4 PRELIMINARY; PRT; 389 AA.
ID Q9WJC4
AC Q9WJC4;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Pre-S and S proteins.
GN S.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=5204397; PubMed=3996597;
RA Bicho V., Drellina D., Pushko P.M., Pumpen P.F., Gren E.;
RT "Subtype ayw variant of hepatitis B virus."
RL FEBS Lett. 185:208-212 (1985).
DR EMBL: X02496; CAB41701.1; .
DR InterPro: IPR000349; HepY1_suffAg.
DR Pfam: PF00695; VMSA; 1.
SQ SSQUNDCR 389 AA; 42745 MW; 6CD83C7CFD2BADF6 CRC64;

Query Match 98.5%; Score 1901; DB 12; Length 389;
Best local Similarity 100.0%; Pred. No. 1.8e-154;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	6	GAGAGGAGTTPPHGGLGWSPOAGILETLPANPPASTNRQSGROPTPLSPPLRTTHPO	65
DB	48	GAGAGGAGTTPPHGGLGWSPOAGILETLPANPPASTNRQSGROPTPLSPPLRTTHPO	107
QY	66	AMQWNSITTHQTLQDPRVGLYFPACGSSSGTVPVTTVSPISSIFSRIGDPALMNI	125
DB	108	AMQWNSITTHQTLQDPRVGLYFPACGSSSGTVPVTTVSPISSIFSRIGDPALMNI	167
QY	126	TSGFLPLVLVQAGFLRLRLITIPQSLDQWTSINFLGTTVCAGQNSQSPSHSPTS	185
DB	168	TSGFLPLVLVQAGFLRLRLITIPQSLDQWTSINFLGTTVCAGQNSQSPSHSPTS	227

QY 186 CPPTCPGYRMCCLRRFIIIFLLICLIFLLVLDYQMLFVCPILPGSSTTSGPCRTC 245
 DB 228 CPPTCPGYRMCCLRRFIIIFLLICLIFLLVLDYQMLFVCPILPGSSTTSGPCRTC 287
 QY 246 TTPAOGTSMYPSCCCTKPSDGNCTCIPISGWAQKFLMENAARFSLVLPVQMFV 305
 DB 288 TTPAOGTSMYPSCCCTKPSDGNCTCIPISGWAQKFLMENAARFSLVLPVQMFV 347
 QY 306 GLSPTVWLSVIMMMYMGPSLSYLSLSPFLPLPIFFCLMVTYI 347
 DB 348 GLSPTVWLSVIMMMYMGPSLSYLSLSPFLPLPIFFCLMVTYI 389

RESULT 2

Q96843 PRELIMINARY; PRT; 389 AA.
 ID 096843
 AC 096843
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Pres. protein.
 GN PRES.
 OS Hepatitis B virus (subtype ayw).
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
 OX NCBI_Taxid=10418;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AYW;
 RX MEDLINE=9732377; PubMed=9188611;
 RA Stoll-Recker S., Repp R., Glebe D., Schaefer S., Kreuder J., Kann M.,
 Lamert P., Gerlich W.H.;
 RT "Transcription of hepatitis B virus in peripheral blood mononuclear
 cells from persistently infected patients."
 RL J. Virol. 71:5399-5407(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AYW;
 RA Repp R.;
 RL Submitted (Aug-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Y07587; CAA68861.1;
 DR InterPro; IPR000349; Hepvir_surfAg.
 DR Pfam; PF00695; VMSA; 1.
 FT CHAIN 164 389 POTENTIAL.
 SQ SEQUENCE 389 AA; 42744 MW; D8DAB1B9B92E57B CRC64;

Query Match

Best Local Similarity 98.4%; Score 1898; DB 12; Length 389;
 Matches 341; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 GAGAFGLGFTPPHGGILGMSPOAGIILETLPANPPASTNRSGROPTPLSPRLRTHPQ 65
 DB 48 GAGAFGLGFTPPHGGILGMSPOAGIILETLPANPPASTNRSGROPTPLSPRLRTHPQ 107
 QY 66 AMQNSTTFHOTLQDPRRGVLPAGSSSGTVNPVPTVSPISISFRIQDPALNMENI 125
 DB 108 AMQNSTTFHOTLQDPRRGVLPAGSSSGTVNPVPTVSPISISFRIQDPALNMENI 167
 QY 126 TSGFLGPLVLVQAGFLLTRILITIPQSLDSWTSINFLGTTVCLGQNSGPTSNHSTPS 185
 DB 168 TSGFLGPLVLVQAGFLLTRILITIPQSLDSWTSINFLGTTVCLGQNSGPTSNHSTPS 227
 QY 186 CPPTCPGYRMCCLRRFIIIFLLICLIFLLVLDYQMLFVCPILPGSSTTSGPCRTC 245
 DB 228 CPPTCPGYRMCCLRRFIIIFLLICLIFLLVLDYQMLFVCPILPGSSTTSGPCRTC 287
 QY 246 TTPAOGTSMYPSCCCTKPSDGNCTCIPISGWAQKFLMENAARFSLVLPVQMFV 305
 DB 288 TTPAOGTSMYPSCCCTKPSDGNCTCIPISGWAQKFLMENAARFSLVLPVQMFV 347
 QY 306 GLSPTVWLSVIMMMYMGPSLSYLSLSPFLPLPIFFCLMVTYI 347
 DB 348 GLSPTVWLSVIMMMYMGPSLSYLSLSPFLPLPIFFCLMVTYI 389

RESULT 3

Q9PMW6 PRELIMINARY; PRT; 389 AA.
 ID 09PMW6
 AC 09PMW6
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Large S protein.
 GN S.
 OS Hepatitis B virus.
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
 OX NCBI_Taxid=10407;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HBV/95-1758, AND HBV/94-11066;
 RX MEDLINE=20109034; PubMed=10640544;
 RA Hannoun C., Horal P., Lindh M.;
 RT "Long-term mutation rates in the hepatitis B virus genome."
 RL J. Gen. Virol. 81:75-83(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HBV/95-1758, AND HBV/94-11066;
 RA Hannoun C., Horal P., Lindh M.;
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF121241; AAF24674.1;
 DR EMBL; AF121240; AAF24667.1;
 DR InterPro; IPR000349; Hepvir_surfAg.
 DR Pfam; PF00695; VMSA; 1.
 SQ SEQUENCE 389 AA; 42814 MW; CD1898538A4A2C0C CRC64;

Query Match

Best Local Similarity 97.7%; Score 1885; DB 12; Length 389;
 Matches 339; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 6 GAGAFGLGFTPPHGGILGMSPOAGIILETLPANPPASTNRSGROPTPLSPRLRTHPQ 65
 DB 48 GAGAFGLGFTPPHGGILGMSPOAGIILETLPANPPASTNRSGROPTPLSPRLRTHPQ 107
 QY 66 AMQNSTTFHOTLQDPRRGVLPAGSSSGTVNPVPTVSPISISFRIQDPALNMENI 125
 DB 108 AMQNSTTFHOTLQDPRRGVLPAGSSSGTVNPVPTVSPISISFRIQDPALNMENI 167
 QY 126 TSGFLGPLVLVQAGFLLTRILITIPQSLDSWTSINFLGTTVCLGQNSGPTSNHSTPS 185
 DB 168 TSGFLGPLVLVQAGFLLTRILITIPQSLDSWTSINFLGTTVCLGQNSGPTSNHSTPS 227
 QY 186 CPPTCPGYRMCCLRRFIIIFLLICLIFLLVLDYQMLFVCPILPGSSTTSGPCRTC 245
 DB 228 CPPTCPGYRMCCLRRFIIIFLLICLIFLLVLDYQMLFVCPILPGSSTTSGPCRTC 287
 QY 246 TTPAOGTSMYPSCCCTKPSDGNCTCIPISGWAQKFLMENAARFSLVLPVQMFV 305
 DB 288 TTPAOGTSMYPSCCCTKPSDGNCTCIPISGWAQKFLMENAARFSLVLPVQMFV 347
 QY 306 GLSPTVWLSVIMMMYMGPSLSYLSLSPFLPLPIFFCLMVTYI 347
 DB 348 GLSPTVWLSVIMMMYMGPSLSYLSLSPFLPLPIFFCLMVTYI 389

RESULT 4

Q67871 PRELIMINARY; PRT; 389 AA.
 ID 067871
 AC 067871
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Pre S1/pre S2/S ORF.
 OS Hepatitis B virus.
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
 OX NCBI_Taxid=10407;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Lai M.E., Mazzoleni A.P., Balestrieri A., Wells A., Portu A.,
 RT "Sequence analysis of HBV genomes isolated from patients with HBeAg
 negative chronic liver disease."
 RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
 DR EMBL; X65257; CA46349.1; -
 DR Interpro; IPR000349; Hepvir_surfac.
 DR Pfam; PF00695; VMSA; 1.
 SQ SEQUENCE 389 AA; 42756 MW; 977025CTDED1A1 CRC64;

Query Match 97.6%; Score 1883; DB 12; Length 389;
 Best Local Similarity 98.8%; Pred. No. 6.2e-153;
 Matches 338; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 GAGAFGIGFTPPHGGILGWSPOAGIILETL PANPPASTNRSGROPTPLSPRLRTHPQ 65
 DB 48 GAGAFGIGFTPPHGGILGWSPOAGIILETL PANPPASTNRSGROPTPLSPRLRTHPQ 107
 QY 66 AMONSTTFHQTLODPRVRLGLYPAGSSSGTVNPVPTVSPSSIFSRIGDPALMMENT 125
 DB 108 AMONSTTFHQTLODPRVRLGLYPAGSSSGTVNPVPTVSPSSIFSRIGDPALMMENT 167
 QY 126 TSGFLGPLVLVLAQGFPLRLILTIPOSLSWMTSLNPLGTTVCLGQNSQSPSHSPTS 185
 DB 168 TSGFLGPLVLVLAQGFPLRLILTIPOSLSWMTSLNPLGTTVCLGQNSQSPSHSPTS 227
 QY 186 CPPTCGYRMWCLRRFIFLFIILLCLIFLVLDYQMLPVCPPLIPSSSTSTGPCRTC 245
 DB 228 CPPTCGYRMWCLRRFIFLFIILLCLIFLVLDYQMLPVCPPLIPSSSTSTGPCRTC 287
 QY 246 TTPAGTSMYPSCCTKPSDGNCTCIPISSNAFGKFLWEMASARPSWLSLVPVQWVF 305
 DB 288 TTPAGTSMYPSCCTKPSDGNCTCIPISSNAFGKFLWEMASARPSWLSLVPVQWVF 347
 QY 306 GLSPVTWLSVIMMMWYMGPSLYSILSPFLPLPIFFCLMVIYI 347
 DB 348 GLSPVTWLSVIMMMWYMGPSLYSILSPFLPLPIFFCLMVIYI 389

RESULT 5

QY 067914 PRELIMINARY; PRT; 389 AA.
 AC 067914
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE Pre-S1 protein.
 OS Hepatitis B virus.
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
 NCBI_TaxID=10407;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97025699; PubMed=8871878;
 RA Alexopoulos A., Karayiannis P., Hadziyannis S.J.;
 RT "Whole genome analysis of hepatitis B virus from 4 cases with
 fulminant hepatitis."
 RL J. Viral Hepat. 3:173-181(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Karayiannis P.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; X97848; CA46425.1; -
 DR Interpro; IPR000349; Hepvir_surfac.
 DR Pfam; PF00695; VMSA; 1.
 SQ SEQUENCE 389 AA; 42717 MW; 495EF28627PBF628 CRC64;

Query Match 97.6%; Score 1883; DB 12; Length 389;
 Best Local Similarity 98.5%; Pred. No. 6.2e-153;
 Matches 337; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 6 GAGAFGIGFTPPHGGILGWSPOAGIILETL PANPPASTNRSGROPTPLSPRLRTHPQ 65
 DB 48 GAGAFGIGFTPPHGGILGWSPOAGIILETL PANPPASTNRSGROPTPLSPRLRTHPQ 107

QY 66 AMONSTTFHQTLODPRVRLGLYPAGSSSGTVNPVPTVSPSSIFSRIGDPALMMENT 125
 DB 108 AMONSTTFHQTLODPRVRLGLYPAGSSSGTVNPVPTVSPSSIFSRIGDPALMMENT 167
 QY 126 TSGFLGPLVLVLAQGFPLRLILTIPOSLSWMTSLNPLGTTVCLGQNSQSPSHSPTS 185
 DB 168 TSGFLGPLVLVLAQGFPLRLILTIPOSLSWMTSLNPLGTTVCLGQNSQSPSHSPTS 227
 QY 186 CPPTCGYRMWCLRRFIFLFIILLCLIFLVLDYQMLPVCPPLIPSSSTSTGPCRTC 245
 DB 228 CPPTCGYRMWCLRRFIFLFIILLCLIFLVLDYQMLPVCPPLIPSSSTSTGPCRTC 287
 QY 246 TTPAGTSMYPSCCTKPSDGNCTCIPISSNAFGKFLWEMASARPSWLSLVPVQWVF 305
 DB 288 TTPAGTSMYPSCCTKPSDGNCTCIPISSNAFGKFLWEMASARPSWLSLVPVQWVF 347
 QY 306 GLSPVTWLSVIMMMWYMGPSLYSILSPFLPLPIFFCLMVIYI 347
 DB 348 GLSPVTWLSVIMMMWYMGPSLYSILSPFLPLPIFFCLMVIYI 389

RESULT 6

QY 067920 PRELIMINARY; PRT; 389 AA.
 AC 067920
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE Pre-S1 protein.
 OS Hepatitis B virus.
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
 NCBI_TaxID=10407;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=PRE-CORE VARIANT;
 RA Alexopoulos A., Karayiannis P., Hadziyannis S.J.;
 RT "Whole genome analysis of hepatitis B virus from 4 cases with
 fulminant hepatitis."
 RL J. Viral Hepat. 3:173-181(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PRE-CORE VARIANT;
 RA Karayiannis P.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; X97849; CA46432.1; -
 DR Interpro; IPR000349; Hepvir_surfac.
 DR Pfam; PF00695; VMSA; 1.
 SQ SEQUENCE 389 AA; 42690 MW; FD05FA0251E162F5 CRC64;

Query Match 97.4%; Score 1879; DB 12; Length 389;
 Best Local Similarity 98.5%; Pred. No. 1.4e-152;
 Matches 337; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 GAGAFGIGFTPPHGGILGWSPOAGIILETL PANPPASTNRSGROPTPLSPRLRTHPQ 65
 DB 48 GAGAFGIGFTPPHGGILGWSPOAGIILETL PANPPASTNRSGROPTPLSPRLRTHPQ 107
 QY 66 AMONSTTFHQTLODPRVRLGLYPAGSSSGTVNPVPTVSPSSIFSRIGDPALMMENT 125
 DB 108 AMONSTTFHQTLODPRVRLGLYPAGSSSGTVNPVPTVSPSSIFSRIGDPALMMENT 167
 QY 126 TSGFLGPLVLVLAQGFPLRLILTIPOSLSWMTSLNPLGTTVCLGQNSQSPSHSPTS 185
 DB 168 TSGFLGPLVLVLAQGFPLRLILTIPOSLSWMTSLNPLGTTVCLGQNSQSPSHSPTS 227
 QY 186 CPPTCGYRMWCLRRFIFLFIILLCLIFLVLDYQMLPVCPPLIPSSSTSTGPCRTC 245
 DB 228 CPPTCGYRMWCLRRFIFLFIILLCLIFLVLDYQMLPVCPPLIPSSSTSTGPCRTC 287
 QY 246 TTPAGTSMYPSCCTKPSDGNCTCIPISSNAFGKFLWEMASARPSWLSLVPVQWVF 305
 DB 288 TTPAGTSMYPSCCTKPSDGNCTCIPISSNAFGKFLWEMASARPSWLSLVPVQWVF 347

QY 306 GLSPTVWLSVIMMMWYMGPSLYSLSPFLPLPIFFCLMVI 347
 DB 348 GLSPTVWLSVIMMMWYMGPSLYSLSPFLPLPIFFCLMVI 389

RESULT 7

QOQMH6 PRELIMINARY; PRT; 389 AA.

AC QOQMH6 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE Pres1/pres2/S envelope.
 GN Hepatitis B virus.
 OS Hepatitis B virus.
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
 OX NCBI_TaxID=10407;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GENOTYPE D;
 RA Okamoto H.;
 RL Submitted (OCT-1999) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GENOTYPE D;
 RX MEDLINE=89010694; PubMed=3171552;
 RA Okamoto H.; Tenda F.; Sakugawa H.; Saetosewignjo R.I.; Imai M.; Miyakawa Y.; Mayumi M.;
 RT "Typing hepatitis B virus by homology in nucleotide sequence:
 RT comparison of surface antigen subtypes."
 RL J. Gen. Virol. 69:2575-2583 (1988).
 DR EMBL; AB033559; BAA85378.1; -
 DR InterPro; IPR000349; Hepvir_surfa.
 DR Pfam; PF00695; vmsa; 1.
 SQ SEQUENCE 389 AA; 42720 MW; 4D60845D294FABE CRC64;

Query Match 97.4%; Score 1878; DB 12; Length 389;

Best Local Similarity 98.5%; Pred. No. 1.7e-152;

Matches 337; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 GAGAFGLGFTPPHGGILGMSPOAGILETLPANPPASTNRSGROPTPLSPRLRTHPQ 65
 DB 48 GAGAFGLGFTPPHGGILGMSPOAGILETLPANPPASTNRSGROPTPLSPRLRTHPQ 107
 QY 66 AMQNSTTFHQTLDPRVRLGYFPAGSSSGTVNPVPTVSPISSIFSRIGDPAIMENI 125
 DB 108 AMQNSTTFHQTLDPRVRLGYFPAGSSSGTVNPVPTVSPISSIFSRIGDPAIMENI 167
 QY 126 TSGFLGLVLYQAGFLLTLTLTIPQSLDSWMTSLNLTGTTVCLGNSGSPSNHSPTS 185
 DB 168 TSGFLGLVLYQAGFLLTLTLTIPQSLDSWMTSLNLTGTTVCLGNSGSPSNHSPTS 227
 QY 186 CPPTCPGRRMCLRRFIIFFILLCIFLLVLDYQMLPVCPLIGSSTTSGPCTC 245
 DB 228 CPPTCPGRRMCLRRFIIFFILLCIFLLVLDYQMLPVCPLIGSSTTSGPCTC 287
 QY 246 TTPAOGTSMYPSCCCTKPSDNCCTCIPSSWAFGKFLMNASRFSWLSLVPVQMFV 305
 DB 288 TTPAOGTSMYPSCCCTKPSDNCCTCIPSSWAFGKFLMNASRFSWLSLVPVQMFV 347
 QY 306 GLSPTVWLSVIMMMWYMGPSLYSLSPFLPLPIFFCLMVI 347
 DB 348 GLSPTVWLSVIMMMWYMGPSLYSLSPFLPLPIFFCLMVI 389

RESULT 8

Q9IF39 PRELIMINARY; PRT; 389 AA.

AC Q9IF39 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)

DE Pres1/pres2/S envelope.

OS Hepatitis B virus.

OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.

OX NCBI_TaxID=10407;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=HBV-D1;

RA Yan L.; Chen J.; Hou J.L.; Wang Z.H.;

RT First Chinese HBV strain: genotype D complete sequence report."

RL Submitted (JUN-2000) to the EMBL/Genbank/DBJ databases.

DR EMBL; AF280817; AAF81608.1; -

DR InterPro; IPR000349; Hepvir_surfa.

DR Pfam; PF00695; vmsa; 1.

SQ SEQUENCE 389 AA; 42784 MW; F53CCFC6A0C1F072 CRC64;

Query Match 97.3%; Score 1876; DB 12; Length 389;

Best Local Similarity 98.5%; Pred. No. 2.5e-152;

Matches 337; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 GAGAFGLGFTPPHGGILGMSPOAGILETLPANPPASTNRSGROPTPLSPRLRTHPQ 65
 DB 48 GAGAFGLGFTPPHGGILGMSPOAGILETLPANPPASTNRSGROPTPLSPRLRTHPQ 107
 QY 66 AMQNSTTFHQTLDPRVRLGYFPAGSSSGTVNPVPTVSPISSIFSRIGDPAIMENI 125
 DB 108 AMQNSTTFHQTLDPRVRLGYFPAGSSSGTVNPVPTVSPISSIFSRIGDPAIMENI 167
 QY 126 TSGFLGLVLYQAGFLLTLTLTIPQSLDSWMTSLNLTGTTVCLGNSGSPSNHSPTS 185
 DB 168 TSGFLGLVLYQAGFLLTLTLTIPQSLDSWMTSLNLTGTTVCLGNSGSPSNHSPTS 227
 QY 186 CPPTCPGRRMCLRRFIIFFILLCIFLLVLDYQMLPVCPLIGSSTTSGPCTC 245
 DB 228 CPPTCPGRRMCLRRFIIFFILLCIFLLVLDYQMLPVCPLIGSSTTSGPCTC 287
 QY 246 TTPAOGTSMYPSCCCTKPSDNCCTCIPSSWAFGKFLMNASRFSWLSLVPVQMFV 305
 DB 288 TTPAOGTSMYPSCCCTKPSDNCCTCIPSSWAFGKFLMNASRFSWLSLVPVQMFV 347
 QY 306 GLSPTVWLSVIMMMWYMGPSLYSLSPFLPLPIFFCLMVI 347
 DB 348 GLSPTVWLSVIMMMWYMGPSLYSLSPFLPLPIFFCLMVI 389

RESULT 9

Q67893 PRELIMINARY; PRT; 389 AA.

AC Q67893 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE Surface protein S.
 GN SURFACE PROTEIN S.
 OS Hepatitis B virus.
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
 OX NCBI_TaxID=10407;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AYM4;
 RA Plutemicak A.;
 RT "Molecular cloning and sequencing of two complete genomes of poliovirus isolates of human Hepatitis B virus."
 RT Submitted (AUG-1994) to the EMBL/Genbank/DBJ databases.
 DR EMBL; Z5716; CAA84788.1; -
 DR InterPro; IPR000349; Hepvir_surfa.
 DR Pfam; PF00695; vmsa; 1.
 DR FTam; PF00695; vmsa; 1.
 SQ SEQUENCE 389 AA; 42762 MW; 5D27F0BA97A825D5 CRC64;

Query Match 97.2%; Score 1875; DB 12; Length 389;

Best Local Similarity 98.2%; Pred. No. 3e-152;

Matches 336; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 6 GAGAFGLGFTPPHGGILGMSPOAGIILETL PANPPASTNRQSGROPTPLSPPLRANTHQ 65
DB 48 GAGAFGLGFTPPHGGILGMSPOAGIILETL PANPPASTNRQSGROPTPLSPPLRANTHQ 107
QY 66 AMONNSTTFHOTLQDPRVRLGYPAGSSSGTVPVPTTSPISISIRIGDPLANMENI 125
DB 108 AMONNSTTFHOTLQDPRVRLGYPAGSSSGTVPVPTTSPISISIRIGDPLANMENI 167
QY 126 TSGFLGLPLVLAQGFPLTRILITIPQSLDSWMTSLNFIAGTTVCLGQNSQSPISNHSPTS 185
DB 168 TSGFLGLPLVLAQGFPLTRILITIPQSLDSWMTSLNFIAGTTVCLGQNSQSPISNHSPTS 227
QY 186 CPPTCGYRMMCLRRFIIIFILLLCLIFLVLDYQGMIPVCPILIPGSSSTTSGPCRTC 245
DB 228 CPPTCGYRMMCLRRFIIIFILLLCLIFLVLDYQGMIPVCPILIPGSSSTTSGPCRTC 287
QY 246 TTPAOGTSMYPSGCCCTKPSDGNCTCIPISSWAFGKFLMWSARFSLVLPVQMFV 305
DB 288 TTPAOGTSMYPSGCCCTKPSDGNCTCIPISSWAFGKFLMWSARFSLVLPVQMFV 347
QY 306 GLSPTVWLSVIMMMYMGPSLSYLSLSPFLPLPIFFCLMAYI 347
DB 348 GLSPTVWLSVIMMMYMGPSLSYLSLSPFLPLPIFFCLMAYI 389

RESULT 10

QY 080X03 PRELIMINARY; PRT; 389 AA.
AC 080X03;
DT 01-JUN-2002 (TREMBLREL. 21, Created)
DT 01-JUN-2002 (TREMBLREL. 21, last sequence update)
DT 01-JUN-2002 (TREMBLREL. 21, last annotation update)
DE Large surface antigen.
GN PREST.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OC NCBI_TaxId=10407;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HBSAG NEGATIVE PATIENT GX27;
RA Kay A.C.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=HBSAG NEGATIVE PATIENT GX27;
RA Jeanter D., Chemin I., Mandrand B., Zoulim F., Trepo C., Kay A.;
RT "Characterization of two Hepatitis B virus populations isolated from a
RL Hepatitis B surface antigen-negative patient.";
DR EMBL; AJ344116; CAC87019.1; -
SQ SEQUENCE 389 AA; 42846 MW; E2A6553DAD135A6 CRC64;

Query Match 97.2%; Score 1875; DB 12; Length 389;
Best Local Similarity 98.5%; Pred. No. 3e-152;
Matches 337; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 GAGAFGLGFTPPHGGILGMSPOAGIILETL PANPPASTNRQSGROPTPLSPPLRANTHQ 65
DB 48 GAGAFGLGFTPPHGGILGMSPOAGIILETL PANPPASTNRQSGROPTPLSPPLRANTHQ 107
QY 66 AMONNSTTFHOTLQDPRVRLGYPAGSSSGTVPVPTTSPISISIRIGDPLANMENI 125
DB 108 AMONNSTTFHOTLQDPRVRLGYPAGSSSGTVPVPTTSPISISIRIGDPLANMENI 167
QY 126 TSGFLGLPLVLAQGFPLTRILITIPQSLDSWMTSLNFIAGTTVCLGQNSQSPISNHSPTS 185
DB 168 TSGFLGLPLVLAQGFPLTRILITIPQSLDSWMTSLNFIAGTTVCLGQNSQSPISNHSPTS 227
QY 186 CPPTCGYRMMCLRRFIIIFILLLCLIFLVLDYQGMIPVCPILIPGSSSTTSGPCRTC 245
DB 228 CPPTCGYRMMCLRRFIIIFILLLCLIFLVLDYQGMIPVCPILIPGSSSTTSGPCRTC 287
QY 246 TTPAOGTSMYPSGCCCTKPSDGNCTCIPISSWAFGKFLMWSARFSLVLPVQMFV 305

DB 288 TTPAOGTSMYPSGCCCTKPSDGNCTCIPISSWAFGKFLMWSARFSLVLPVQMFV 347
QY 306 GLSPTVWLSVIMMMYMGPSLSYLSLSPFLPLPIFFCLMAYI 347
DB 348 GLSPTVWLSVIMMMYMGPSLSYLSLSPFLPLPIFFCLMAYI 389

RESULT 11

QY 080X09 PRELIMINARY; PRT; 389 AA.
AC 080X09;
DT 01-JUN-2002 (TREMBLREL. 21, Created)
DT 01-JUN-2002 (TREMBLREL. 21, last sequence update)
DT 01-JUN-2002 (TREMBLREL. 21, last annotation update)
DE Large surface antigen.
GN PREST.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OC NCBI_TaxId=10407;
RN [1]
RP SEQUENCE FROM N.A.
RC Kay A.C.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC Jeanter D., Chemin I., Mandrand B., Zoulim F., Trepo C., Kay A.;
RT "Characterization of two Hepatitis B virus populations isolated from a
RL Hepatitis B surface antigen-negative patient.";
DR EMBL; AJ344117; CAC87013.1; -
SQ SEQUENCE 389 AA; 42764 MW; 9AC5A4D46B73632 CRC64;

Query Match 97.2%; Score 1875; DB 12; Length 389;
Best Local Similarity 98.2%; Pred. No. 3e-152;
Matches 336; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 6 GAGAFGLGFTPPHGGILGMSPOAGIILETL PANPPASTNRQSGROPTPLSPPLRANTHQ 65
DB 48 GAGAFGLGFTPPHGGILGMSPOAGIILETL PANPPASTNRQSGROPTPLSPPLRANTHQ 107
QY 66 AMONNSTTFHOTLQDPRVRLGYPAGSSSGTVPVPTTSPISISIRIGDPLANMENI 125
DB 108 AMONNSTTFHOTLQDPRVRLGYPAGSSSGTVPVPTTSPISISIRIGDPLANMENI 167
QY 126 TSGFLGLPLVLAQGFPLTRILITIPQSLDSWMTSLNFIAGTTVCLGQNSQSPISNHSPTS 185
DB 168 TSGFLGLPLVLAQGFPLTRILITIPQSLDSWMTSLNFIAGTTVCLGQNSQSPISNHSPTS 227
QY 186 CPPTCGYRMMCLRRFIIIFILLLCLIFLVLDYQGMIPVCPILIPGSSSTTSGPCRTC 245
DB 228 CPPTCGYRMMCLRRFIIIFILLLCLIFLVLDYQGMIPVCPILIPGSSSTTSGPCRTC 287
QY 246 TTPAOGTSMYPSGCCCTKPSDGNCTCIPISSWAFGKFLMWSARFSLVLPVQMFV 305
DB 288 TTPAOGTSMYPSGCCCTKPSDGNCTCIPISSWAFGKFLMWSARFSLVLPVQMFV 347
QY 306 GLSPTVWLSVIMMMYMGPSLSYLSLSPFLPLPIFFCLMAYI 347
DB 348 GLSPTVWLSVIMMMYMGPSLSYLSLSPFLPLPIFFCLMAYI 389

RESULT 12

QY 090AF2 PRELIMINARY; PRT; 389 AA.
AC 090AF2;
DT 01-MAY-2000 (TREMBLREL. 13, Created)
DT 01-MAY-2000 (TREMBLREL. 13, last sequence update)
DT 01-MAY-2000 (TREMBLREL. 20, last annotation update)
DE Large s protein.
GN S.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.

OX NCBI_TaxID=10407;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HBV/1218;
 RX MEDLINE=20109034; PubMed=10640544;
 RA Hannou C., Horal P., Lindh M.;
 RT "Long-term mutation rates in the hepatitis B virus genome";
 RL J. Gen. Virol. 81:75-83(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HBV/1218;
 RA Hannou C., Horal P., Lindh M.;
 RL Submitted (JAN-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL, AF121242; AAF24681.1; -;
 DR InterPro; IPR000349; Hepvlt_surflag.
 DR Pfam; PF00695; VMSA; 1.
 SQ SEQUENCE 389 AA; 42910 MW; BB6541E18D39357E CRC64;
 Query Match 97.1%; Score 1874; DB 12; Length 389;
 Best Local Similarity 98.5%; Pred. No. 3.7e-152;
 Matches 337; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 6 GAGAFGLGFTPHGGLGMSPOAGILETLPANPPASTNRSGRPPLSPPLANTHQ 65
 DB 48 GAGAFGLGFTPHGGLGMSPOAGILETLPTNPPASTNRSGRPPLSPPLANTHQ 107
 QY 66 ANOMNSTTHQTLQDRVAGLVFPAGSSSGTVNPPVPTVSPISSIFSRIGDPALMNI 125
 DB 108 ANOMNSTTHQTLQDRVAGLVFPAGSSSGTVNPPVPTVSPISSIFSRIGDPALMNI 167
 QY 126 TSGFLPLVLQAGFPLRLILTPQSLDSWMTSLNPLGTTVCIGQNSQSPSHSPTS 185
 DB 168 TSGFLPLVLQAGFPLRLILTPQSLDSWMTSLNPLGTTVCIGQNSQSPSHSPTS 227
 QY 186 CPPTCGYRMNCLRRPILFLFILLCLIFLVLDYQGLPVCPPLPGSSTTSGPCRTC 245
 DB 228 CPPTCGYRMNCLRRPILFLFILLCLIFLVLDYQGLPVCPPLPGSSTTSGPCRTC 287
 QY 246 TTPAOGTSMYPSCCCTKPSDGNCTCIPISMSAFGKFLMWSARFSWLSLVPVQMFV 305
 DB 288 TTPAOGTSMYPSCCCTKPSDGNCTCIPISMSAFGKFLMWSARFSWLSLVPVQMFV 347
 QY 306 GLSPTVMSVIMMMWYGPSTLSLSPFLPLPIFCLMAYI 347
 DB 348 GLSPTVMSVIMMMWYGPSTLSLSPFLPLPIFCLMAYI 389
 RESULT 13
 ID Q9QAP9 PRELIMINARY; PRT; 389 AA.
 AC Q9QAP9;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Large S protein.
 GN S.
 OS Hepatitis B virus.
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
 OC NCBI_TaxID=10407;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HBV/1303;
 RX MEDLINE=20109034; PubMed=10640544;
 RA Hannou C., Horal P., Lindh M.;
 RT "Long-term mutation rates in the hepatitis B virus genome";
 RL J. Gen. Virol. 81:75-83(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HBV/1303;
 RA Hannou C., Horal P., Lindh M.;
 RL Submitted (JAN-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL, AF121239; AAF24661.1; -;
 DR InterPro; IPR000349; Hepvlt_surflag.

DR Pfam; PF00695; VMSA; 1.
 SQ SEQUENCE 389 AA; 42890 MW; 3D971AB01FD77FBC CRC64;
 Query Match 96.6%; Score 1864; DB 12; Length 389;
 Best Local Similarity 98.5%; Pred. No. 2.6e-151;
 Matches 337; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 6 GAGAFGLGFTPHGGLGMSPOAGILETLPANPPASTNRSGRPPLSPPLANTHQ 65
 DB 48 GAGAFGLGFTPHGGLGMSPOAGILETLPTNPPASTNRSGRPPLSPPLANTHQ 107
 QY 66 ANOMNSTTHQTLQDRVAGLVFPAGSSSGTVNPPVPTVSPISSIFSRIGDPALMNI 125
 DB 108 ANOMNSTTHQTLQDRVAGLVFPAGSSSGTVNPPVPTVSPISSIFSRIGDPALMNI 167
 QY 126 TSGFLPLVLQAGFPLRLILTPQSLDSWMTSLNPLGTTVCIGQNSQSPSHSPTS 185
 DB 168 TSGFLPLVLQAGFPLRLILTPQSLDSWMTSLNPLGTTVCIGQNSQSPSHSPTS 227
 QY 186 CPPTCGYRMNCLRRPILFLFILLCLIFLVLDYQGLPVCPPLPGSSTTSGPCRTC 245
 DB 228 CPPTCGYRMNCLRRPILFLFILLCLIFLVLDYQGLPVCPPLPGSSTTSGPCRTC 287
 QY 246 TTPAOGTSMYPSCCCTKPSDGNCTCIPISMSAFGKFLMWSARFSWLSLVPVQMFV 305
 DB 288 TTPAOGTSMYPSCCCTKPSDGNCTCIPISMSAFGKFLMWSARFSWLSLVPVQMFV 347
 QY 306 GLSPTVMSVIMMMWYGPSTLSLSPFLPLPIFCLMAYI 347
 DB 348 GLSPTVMSVIMMMWYGPSTLSLSPFLPLPIFCLMAYI 389
 RESULT 14
 ID Q9QNS1 PRELIMINARY; PRT; 389 AA.
 AC Q9QNS1;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Pre-S1 protein precursor.
 GN S.
 OS Hepatitis B virus.
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
 OC NCBI_TaxID=10407;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AYW2;
 RX MEDLINE=20006001; PubMed=10534721;
 RA Gerner P., Lausch E., Friedt M., Tratzmuller R., Spangenberg C.,
 RA Wirth S.;
 RT "Hepatitis B virus core promoter mutations in children with multiple
 RT anti-HBe/HBeAg reactivations result in enhanced promoter activity";
 RL J. Med. Virol. 59:415-423(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AYW2;
 RA Gerner P., Lausch E., Friedt M., Tratzmuller R., Spangenberg C.,
 RA Wirth S.;
 RT Submitted (MAY-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL, AF151735; AAD41361.1; -;
 DR InterPro; IPR000349; Hepvlt_surflag.
 DR Pfam; PF00695; VMSA; 1.
 KM Signal.
 FT SIGNAL. 1 54 POTENTIAL.
 FT CHAIN 164 389 S PROTEIN.
 SQ SEQUENCE 389 AA; 42779 MW; 30511B862904BE7C CRC64;
 Query Match 96.6%; Score 1864; DB 12; Length 389;
 Best Local Similarity 98.0%; Pred. No. 2.6e-151;
 Matches 335; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 QY 6 GAGAFGLGFTPHGGLGMSPOAGILETLPANPPASTNRSGRPPLSPPLANTHQ 65


```

Db 48 GAGAGGAGTTPPHGGLGMSPOAGILHTVPANPPASTNRQSGROPTPLSPPLRNTHPQ 107
QY 66 AMONNSTTFHOTLDDPRVRGLYPPAGSSSGTVPNPTTVPSSISIFSRIGDPALNMENI 125
Db 108 AMONNSTTFHOTLDDPRVRGLYPPAGSSSGTVPNPTTVPSSISIFSRIGDPALNMENI 167
QY 126 TSGFLGPIVLVQAQGFILTRILITIPQSLDSWTSINFLGCTTVCCLGONSQSPSNHSPTS 185
Db 168 TSGFLGPIVLVQAQGFILTRILITIPQSLDSWTSINFLGCTTVCCLGONSQSPSNHSPTS 227
QY 186 CPPTCPGRWMCRLRRFIFLFIILLCILFILVLDYQGMPLVPCPLIPGSSSTTSGPCRTC 245
Db 228 CPPTCPGRWMCRLRRFIFLFIILLCILFILVLDYQGMPLVPCPLIPGSSSTTSGPCRTC 287
QY 246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLMEMASARFSWLSLVPVQMFV 305
Db 288 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLMEMASARFSWLSLVPVQMFV 347
QY 306 GLSPTVWLSVIMMMWYGPISLYSILSPFLPLPIFFCLMWYI 347
Db 348 GLSPTVWLSVIMMMWYGPISLYSILSPFLPLPIFFCLMWYI 389

```

RESULT 15

```

ID 092921 PRELIMINARY; PRT; 389 AA.
AC 092921;
DT 01-NOV-1998 (TEMBLrel. 08, Created)
DT 01-NOV-1998 (TEMBLrel. 08, Last sequence update)
DT 01-MAR-2002 (TEMBLrel. 20, Last annotation update)
DE Pre-S1 protein.
GN S.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_Taxid=10407;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1/91;
RX MEDLINE=98240943; PubMed=9581787;
RA Gunther S., Wilma P., Meisel H., Will H.;
RT "Analysis of hepatitis B virus populations in an interferon-alpha-
RT treated patient reveals predominant mutations in the C-gene and
RT changing e-antigenicity."
RL Virology 244:146-160(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=1/91;
RA Gunther S., Wilma P., Meisel H., Will H.;
RT Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF043594; AAC40811.1;
DR InterPro; IPR000349; Hepvit_surfa9.
DR Pfam; PF00695; VMSA; 1.
FT CHAIN 164 389 S PROTEIN.
SQ SEQUENCE 389 AA; 42795 MW; 3369787D78F188B9 CRC64;

```

Query Match 96.5%; Score 1861; DB 12; Length 389;
 Best Local Similarity 97.4%; Pred. No. 4,7e-151;
 Matches 333; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

```

QY 6 GAGAGGAGTTPPHGGLGMSPOAGILHTVPANPPASTNRQSGROPTPLSPPLRNTHPQ 65
Db 48 GAGAGGAGTTPPHGGLGMSPOAGILHTVPANPPASTNRQSGROPTPLSPPLRNTHPQ 107
QY 66 AMONNSTTFHOTLDDPRVRGLYPPAGSSSGTVPNPTTVPSSISIFSRIGDPALNMENI 125
Db 108 AMONNSTTFHOTLDDPRVRGLYPPAGSSSGTVPNPTTVPSSISIFSRIGDPALNMENI 167
QY 126 TSGFLGPIVLVQAQGFILTRILITIPQSLDSWTSINFLGCTTVCCLGONSQSPSNHSPTS 185
Db 168 TSGFLGPIVLVQAQGFILTRILITIPQSLDSWTSINFLGCTTVCCLGONSQSPSNHSPTS 227
QY 186 CPPTCPGRWMCRLRRFIFLFIILLCILFILVLDYQGMPLVPCPLIPGSSSTTSGPCRTC 245

```

```

Db 228 CPPTCPGRWMCRLRRFIFLFIILLCILFILVLDYQGMPLVPCPLIPGSSSTTSVGPCRTC 287
QY 246 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLMEMASARFSWLSLVPVQMFV 305
Db 288 TTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLMEMASARFSWLSLVPVQMFV 347
QY 306 GLSPTVWLSVIMMMWYGPISLYSILSPFLPLPIFFCLMWYI 347
Db 348 GLSPTVWLSVIMMMWYGPISLYSILSPFLPLPIFFCLMWYI 389

```

Search completed: February 3, 2003, 09:25:43
 Job time : 48.8821 secs

OY 186 CPPTCPGYRMCLRRFIIIFLLILLCLIFLLVLDYGMFLVCPILIPSSSTTSTGPCRTC 245
 DB 228 CPPTCPGYRMCLRRFIIIFLLILLCLIFLLVLDYGMFLVCPILIPSSSTTSTGPCRTC 287
 OY 246 TTPAOGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLMWSARFWSLILVFPVQMFV 305
 DB 288 TTPAOGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLMWSARFWSLILVFPVQMFV 347
 OY 306 GLSPFTWLSVITMMWYMGPSLSYLSLSPFLPLPIFFCLMWTYI 347
 DB 348 GLSPFTWLSVITMMWYMGPSLSYLSLSPFLPLPIFFCLMWTYI 389
 RESULT 2
 VMSA_HPBVZ STANDARD; PRT; 389 AA.
 ID VMSA_HPBVZ
 AC P03139; 1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Major surface antigen precursor.
 GN S.
 OS Hepatitis B virus (subtype ayw).
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
 OX NCBI_TaxId=10418;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=81012091; PubMed=399329;
 RA Galibert F., Mandart E., Fricouat F., Tiollais P., Charnay P.;
 RT "Nucleotide sequence of the hepatitis B virus genome (subtype ayw)
 RT cloned in E. coli.";
 RT Nature 281:646-650 (1979).
 RN [2]
 RP SEQUENCE FROM N.A. (CLONE PHB320).
 RX MEDLINE=85204397; PubMed=396597;
 RA Bichko V., Puhko P., Dreilina D., Pumpen P., Gren E.Y.;
 RT "Subtype ayw variant of hepatitis B virus. DNA primary structure
 RT analysis.";
 RT FEBS Lett. 185:208-212 (1985).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Pumpen P., Kozlovskaya T.M., Borisova G.L., Bunchko V.V.,
 RA Dushler A.V., Kalle Y.V., Pudova N.V., Gren E.Y., Kravtynya V.Y.,
 RA Kukain R.A.;
 RT "Synthesis of the surface antigen of hepatitis B virus in Escherichia
 RT coli.";
 RT Dokl. Biochem. 271:246-249 (1984).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; V01460; -; NOT ANNOTATED_CDS.
 DR EMBL; X02496; CA26324.1; ALT_INIT.
 DR EMBL; M12393; AAA45496.1; ALT_INIT.
 DR PIR; A03703; SAVLAH.
 DR PIR; A03704; SAVLBH.
 DR InterPro; IPR000349; Hepvlt_surfAg.
 DR Pfam; PF00695; VMSA; 1.
 KW Antigen.
 FT PROPEP 1 163
 FT CHAIN 164 389
 FT CARBOHYD 4 4
 FT CARBOHYD 112 112
 FT CARBOHYD 166 166
 FT CARBOHYD 309 309
 FT CARBOHYD 75 75
 FT CONFLICT 147 147
 MAJOR SURFACE ANTIGEN. (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 O -> B (IN REF. 2).
 A -> S (IN REF. 2).

FT CONFLICT 150 150 L -> I (IN REF. 2).
 FT CONFLICT 288 290 MTT -> TTP (IN REF. 2 AND 3).
 SO SEQUENCE 389 AA; 42766 MW; 6DC9B682DA694F63 CRC64;
 Query Match 96.8%; Score 1868; DB 1; Length 389;
 Best Local Similarity 98.2%; Pred. No. 2, 3e-125;
 Matches 336; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 OY 6 GAGAFGLGFTPPHGGILGMSPOAGILETLPANPPASTNRSGROPTPLSPPLANTHQ 65
 DB 48 GAGAFGLGFTPPHGGILGMSPOAGILQTLPANPPASTNRSGROPTPLSPPLANTHQ 107
 OY 66 AMQNSTTTFHQTLQDPVRGLYFPAGSSSGTYNPPVPTTYSPLSSIFSRIGDPLAMNENT 125
 DB 108 AMQNSTTTFHQTLQDPVRGLYFPAGSSSGTYNPPVPTTYSPLSSIFSRIGDPLAMNENT 167
 OY 126 TSGFLGSLVLOAGFLLTFLTLTIPSLDSMWTSLNPLFGTTCVCLQNSQSPSNHSPTS 185
 DB 168 TSGFLGSLVLOAGFLLTFLTLTIPSLDSMWTSLNPLFGTTCVCLQNSQSPSNHSPTS 227
 OY 186 CPPTCPGYRMCLRRFIIIFLLILLCLIFLLVLDYGMFLVCPILIPSSSTTSTGPCRTC 245
 DB 228 CPPTCPGYRMCLRRFIIIFLLILLCLIFLLVLDYGMFLVCPILIPSSSTTSTGPCRTC 287
 OY 246 TTPAOGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLMWSARFWSLILVFPVQMFV 305
 DB 288 TTPAOGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLMWSARFWSLILVFPVQMFV 347
 OY 306 GLSPFTWLSVITMMWYMGPSLSYLSLSPFLPLPIFFCLMWTYI 347
 DB 348 GLSPFTWLSVITMMWYMGPSLSYLSLSPFLPLPIFFCLMWTYI 389

RESULT 3
 VMSA_HPBVZ STANDARD; PRT; 389 AA.
 ID VMSA_HPBVZ
 AC P03139;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 01-OCT-1989 (Rel. 12, Last annotation update)
 DE Major surface antigen precursor.
 GN S.
 OS Hepatitis B virus (subtype ayw).
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
 OX NCBI_TaxId=10418;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=81012115; PubMed=399329;
 RA Paek M., Goto T., Gilbert W., Zink B., Schaller H., Mackay P.,
 RA Leadbetter G., Murray K.;
 RT "Hepatitis B virus genes and their expression in E. coli.";
 RT Nature 282:575-579 (1979).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; J02202; AAA45487.1; ALT_INIT.
 DR PIR; A93217; SAVLAJ.
 DR InterPro; IPR000349; Hepvlt_surfAg.
 DR Pfam; PF00695; VMSA; 1.
 KW Antigen.
 FT PROPEP 1 163
 FT CHAIN 164 389
 FT CARBOHYD 4 4
 FT CARBOHYD 112 112
 FT CARBOHYD 166 166
 FT CARBOHYD 309 309
 FT CARBOHYD 389 389
 MAJOR SURFACE ANTIGEN. (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SO SEQUENCE 389 AA; 42801 MW; 1069BFB64679A669 CRC64;

Query Match
Best Local Similarity 96.04; Score 1851; DB 1; Length 389;
Matches 332; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 6 GAGAFGLGFTPPHGLGMSPOAGIILETLPANPPASTNROSGROPTPLSPRLRTHPQ 65
DB 48 GAGAFGLGFTPPHGLGMSPOAGIILETLPANPPASTNROSGROPTPLSPRLRTHPQ 107

QY 66 AMONSTTFHQTLDPRVRGLYPAGGSSGTVNPPVTTPSPISISRSRIGDPLANNMENT 125
DB 108 AMONSTTFHQTLDPRVRGLYPAGGSSGTVNPPVTTPSPISISRSRIGDPLANNMENT 167

QY 126 TSGFLGFLVLQAGFFLLTRILITIPQSLDSWMTSLNFGSTTVCLGQNSQSPISNHSPTS 185
DB 168 TSGFLGFLVLQAGFFLLTRILITIPQSLDSWMTSLNFGSTTVCLGQNSQSPISNHSPTS 227

QY 186 CPPTCGYRMWCLRRFTIFILLCLIFLLVLDYQMLPVCPLIPGSSSTTSSTGCRCTC 245
DB 228 CPPTCGYRMWCLRRFTIFILLCLIFLLVLDYQMLPVCPLIPGSSSTTSSTGCRCTC 287

QY 246 TTPAOGTSMWPCSCCTKPSDGNCTCIPBSSNAFGKFLWEMASARFSLVLPVQWVF 305
DB 288 TTPAOGTSMWPCSCCTKPSDGNCTCIPBSSNAFGKFLWEMASARFSLVLPVQWVF 347

QY 306 GLSPTVWLSVIMMMWYGPGLYSILSPFLPLIPFLCLAWYI 347
DB 348 GLSPTVWLSVIMMMWYGPGLYSILSPFLPLIPFLCLAWYI 389

RESULT 4
VMSA_HPBVR STANDARD; PRT; 400 AA.

ID VMSA_HPBVR
AC P03140;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Major surface antigen precursor.
GN S.
OS Hepatitis B virus (subtype adr).
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=106820;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83168919; PubMed=6300776;
RA Ono Y., Onda H., Sasaki R., Igatahshi K., Sugino Y., Nishiohka K.;
RT "The complete nucleotide sequences of the cloned hepatitis B virus
RL DNA: subtype adr and adm."
NU Nucleic Acids Res. 11:1747-1757(1983).
RN [2]
RP SEQUENCE OF 176-350 FROM N.A.
RX MEDLINE=85107103; PubMed=3968537;
RA Takeshima H., Inokoshi J., Namiaki M., Shimada J., Omura S.;
RT "Structural analysis of the gene coding for hepatitis B virus surface
RL antigen and its product."
J. Gen. Virol. 66:195-200(1985).
CC CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: V00867; CAA24234.1; ALT_INIT.
DR PIR: A03705; SAVIA.
DR InterPro: IPR000349; Hepv1r_surfAg.
DR Pfam: PF00695; VMSA; 1.
DR Antigen.
KW Antigen.
FT PROPEP 1 174
FT CHAIN 175 400 MAJOR SURFACE ANTIGEN.
FT CARBOHYD 15 15 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 123 123 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 320 320 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 400 AA; 43685 MW; 16DB0B1566C697A CRC64;

Query Match
Best Local Similarity 92.18; Score 1771; DB 1; Length 400;
Matches 315; Conservative 10; Mismatches 17; Indels 0; Gaps 0;

QY 6 GAGAFGLGFTPPHGLGMSPOAGIILETLPANPPASTNROSGROPTPLSPRLRTHPQ 65
DB 59 GAGAFGLGFTPPHGLGMSPOAGIILETLPANPPASTNROSGROPTPLSPRLRTHPQ 118

QY 66 AMONSTTFHQTLDPRVRGLYPAGGSSGTVNPPVTTPSPISISRSRIGDPLANNMENT 125
DB 119 AMONSTTFHQTLDPRVRGLYPAGGSSGTVNPPVTTPSPISISRSRIGDPLANNMENT 178

QY 126 TSGFLGFLVLQAGFFLLTRILITIPQSLDSWMTSLNFGSTTVCLGQNSQSPISNHSPTS 185
DB 179 TSGFLGFLVLQAGFFLLTRILITIPQSLDSWMTSLNFGSTTVCLGQNSQSPISNHSPTS 238

QY 186 CPPTCGYRMWCLRRFTIFILLCLIFLLVLDYQMLPVCPLIPGSSSTTSSTGCRCTC 245
DB 239 CPPTCGYRMWCLRRFTIFILLCLIFLLVLDYQMLPVCPLIPGSSSTTSSTGCRCTC 298

QY 246 TTPAOGTSMWPCSCCTKPSDGNCTCIPBSSNAFGKFLWEMASARFSLVLPVQWVF 305
DB 299 TTPAOGTSMWPCSCCTKPSDGNCTCIPBSSNAFGKFLWEMASARFSLVLPVQWVF 358

QY 306 GLSPTVWLSVIMMMWYGPGLYSILSPFLPLIPFLCLAWYI 347
DB 359 GLSPTVWLSVIMMMWYGPGLYSILSPFLPLIPFLCLAWYI 400

RESULT 5
VMSA_HPBVL STANDARD; PRT; 389 AA.

ID VMSA_HPBVL
AC P12911;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-APR-1990 (Rel. 14, Last annotation update)
DE Major surface antigen precursor.
GN S.
OS Hepatitis B virus (strain 1sh / chimpanzee isolate).
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10414;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88258473; PubMed=2838576;
RA Vaudin M., Wolstenholme A.J., Tsiquaye K.N., Zuckerman A.J.;
RT "The complete nucleotide sequence of the genome of a hepatitis B
RL virus isolated from a naturally infected chimpanzee."
J. Gen. Virol. 69:1383-1389(1988).
CC CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: D00220; BAA00159.1; -.
DR PIR: C28885; SAVICP.
DR InterPro: IPR000349; Hepv1r_surfAg.
DR Pfam: PF00695; VMSA; 1.
DR Antigen.
KW Antigen.
FT PROPEP 1 163
FT CHAIN 164 389
FT CARBOHYD 4 4
FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 309 309 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 389 AA; 42539 MW; 87D1P93382A10BDD CRC64;

Query Match 91.0%; Score 1755; DB 1; Length 389;
Best Local Similarity 91.5%; Pred. No. 2.3e-117;
Matches 313; Conservative 9; Mismatches 20; Indels 0; Gaps 0;

QY 6 GAGAGCAGTTPPHGGLGMSPOAGILETLPANPPASTNOSGROPTPLSPPLRNTHPQ 65
DB 48 GAGAGGPGTTPPHGGLGMSPOAGILETLPANPPASTNOSGROPTPLSPPLRNTHPQ 107
QY 66 AMQNSTTHQTLQDPRVAGLYFPAGSSSGTVNVPPTVSPISISIFRIDPPALNMENI 125
DB 108 AMQNSTTHQTLQDPRVAGLYFPAGSSSGTVNVPPTVSPISISIFRIDPPALNMENI 167
QY 126 TSGFLGPLVLVLAQGFPLRLITLIPQSLDSWMTSLNPLGAGAVCLGQNSQSPFTHSPTS 185
DB 168 TSGFLGPLVLVLAQGFPLRLITLIPQSLDSWMTSLNPLGAGAVCLGQNSQSPFTHSPTS 227
QY 186 CPPTCGYRMCCLRRPILFLFILLCLFLVLDYQGLPVCPLIPGSSSTGRCRTG 245
DB 228 CPPTCGYRMCCLRRPILFLFILLCLFLVLDYQGLPVCPLIPGSSSTGRCRTG 287
QY 246 TTPAOGTSMYPSCCCTKPSDGNCTCIPISNAFGKFLMNASAFSWLSLVPVQWFW 305
DB 288 TTPAOGTSMYPSCCCTKPSDGNCTCIPISNAFGKFLMNASAFSWLSLVPVQWFW 347
QY 306 GLSPTVWLSIWMWYMGPSLYSILSPFLPLPFPCLMWYI 347
DB 348 GLSPTVWLSIWMWYMGPSLYSILSPFLPLPFPCLMWYI 389

RESULT 6
VMSA_HPBVP STANDARD; PRT; 400 AA.

ID VMSA_HPBVP 002317;
AC 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE Major surface antigen precursor.
GN 5.
OS Hepatitis B virus (subtype adr / strain Philippines/PFDM294).
OS Viruses; Retroviral viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=31514;
[1]
RP SEQUENCE FROM N.A.
RA Betaccio R.C., Chavez C.C., Okamoto H., Lingao A.L., Reyes M.T.,
RA Domingo E., Mayumi M.,
RT "Nucleotide sequence of a hepatitis B virus genome of subtype adr
RT isolated from a Philippine: comparison with the reported three genomes
RT of the same subtype."
RL J. Gastroenterol. Hepatol. 3:215-222(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M57663; AA69680.1; -
DR InterPro; IPR000349; Hepvlt_surfAg.
DR Pfam; PF00695; VMSA; 1.
KW Antigen.

FT PROPEP 1 174 MAJOR SURFACE ANTIGEN. (POTENTIAL).
FT CHAIN 175 400
FT CARBOHYD 15 15 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 123 123 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 320 320 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 400 AA; 43740 MW; 7EBD53A004136BD2 CRC64;

Query Match 91.0%; Score 1755; DB 1; Length 400;
Best Local Similarity 90.9%; Pred. No. 2.4e-117;
Matches 311; Conservative 13; Mismatches 18; Indels 0; Gaps 0;

QY 6 GAGAGCAGTTPPHGGLGMSPOAGILETLPANPPASTNOSGROPTPLSPPLRNTHPQ 65
DB 59 GAGAGGPGTTPPHGGLGMSPOAGILETLPANPPASTNOSGROPTPLSPPLRNTHPQ 118
QY 66 AMQNSTTHQTLQDPRVAGLYFPAGSSSGTVNVPPTVSPISISIFRIDPPALNMENI 125
DB 119 AMQNSTTHQTLQDPRVAGLYFPAGSSSGTVNVPPTVSPISISIFRIDPPALNMENI 178
QY 126 TSGFLGPLVLVLAQGFPLRLITLIPQSLDSWMTSLNPLGAGAVCLGQNSQSPFTHSPTS 185
DB 179 TSGFLGPLVLVLAQGFPLRLITLIPQSLDSWMTSLNPLGAGAVCLGQNSQSPFTHSPTS 238
QY 186 CPPTCGYRMCCLRRPILFLFILLCLFLVLDYQGLPVCPLIPGSSSTGRCRTG 245
DB 239 CPPTCGYRMCCLRRPILFLFILLCLFLVLDYQGLPVCPLIPGSSSTGRCRTG 298
QY 246 TTPAOGTSMYPSCCCTKPSDGNCTCIPISNAFGKFLMNASAFSWLSLVPVQWFW 305
DB 299 TTPAOGTSMYPSCCCTKPSDGNCTCIPISNAFGKFLMNASAFSWLSLVPVQWFW 358
QY 306 GLSPTVWLSIWMWYMGPSLYSILSPFLPLPFPCLMWYI 347
DB 359 GLSPTVWLSIWMWYMGPSLYSILSPFLPLPFPCLMWYI 400

RESULT 7
VMSA_HPBV4 STANDARD; PRT; 400 AA.

ID VMSA_HPBV4 012934; 067860;
AC 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-OCT-1989 (Rel. 12, Last annotation update)
DE Major surface antigen precursor.
GN 5.
OS Hepatitis B virus (subtype adr4).
OS Viruses; Retroviral viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10409;
[1]
RP SEQUENCE FROM N.A.
RA MEDLINE=83246570; PubMed=3306594;
RA Fujiyama A., Miyanozawa A., Nozaki C., Yoneyama T., Ohtomo N.,
RA Matsubara K.;
RT "Cloning and structural analyses of hepatitis B virus DNAs, subtype
RT adr."
RL Nucleic Acids Res. 11:4601-4610(1983).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X01587; CA25747.1; ALT INIT.
DR EMBL; X01587; CA25743.1; ALT INIT.
DR InterPro; IPR000349; Hepvlt_surfAg.
DR Pfam; PF00695; VMSA; 1.
KW Antigen.

FT PROPEP 1 174 MAJOR SURFACE ANTIGEN. (POTENTIAL).
FT CHAIN 175 400
FT CARBOHYD 15 15 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 123 123 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 320 320 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 400 AA; 43542 MW; 01053BBD24053F9 CRC64;

Query Match 90.5%; Score 1745; DB 1; Length 400;
Best Local Similarity 90.9%; Pred. No. 1.2e-116;

Matches 311; Conservative 11; Mismatches 20; Indels 0; Gaps 0;

QY 6 GAGAFGLGFTPPHGGILGMSPOAGILFTLPANPPASTNRSGROPTPLSPPLANTHPQ 65
 DB 59 GAGAFGGFTPPHGGILGMSPOAGILFTLPANPPASTNRSGROPTPLSPPLANTHPQ 118
 QY 66 AMONNSTTFHOTLADPRVRLYFPAGSSSGTVDVPTVSPISISIFSRIGDPALMMENT 125
 DB 119 AMONNSTTFHOTLADPRVRLYFPAGSSSGTVDVPTVSPISISISIRGDPALMMENT 178
 QY 126 TSGFLGPLLVLQAGFFLLTRILITIPQSLDSWMTSLNFIAGTTCVCLGONSOSPNSHPTS 185
 DB 179 TSGFLGPLLVLQAGFFLLTRILITIPQSLDSWMTSLNFIAGTTCVCLGONSOSPNSHPTS 238
 QY 186 CPPTCGYRMWMLRRRIIFLLCLILFLVLDYQGLMPCVCLIPGSSSTTSGPCRTC 245
 DB 239 CPPTCGYRMWMLRRRIIFLLCLILFLVLDYQGLMPCVCLIPGSSSTTSGPCRTC 298
 QY 246 TTPAOGTSMYPSCCCTKPSDGNCTCIPIPSSWAFKFLMNASARFWSLTVFVQWFW 305
 DB 299 TTPAOGTSMYPSCCCTKPSDGNCTCIPIPSSWAFKFLMNASARFWSLTVFVQWFW 358
 QY 306 GLSPTVMTSLVIMMMWYMGPSLYSLSPPLPLPIFCLMVTYI 347
 DB 359 GLSPTVMTSLVIMMMWYMGPSLYSLSPPLPLPIFCLMVTYI 400

RESULT 8
 VMSA_HPBVJ STANDARD; PRT; 389 AA.

ID VMSA_HPBVJ
 AC P17398;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 01-AUG-1992 (Rel. 23, Last annotation update)
 DE Major surface antigen precursor.
 GN S.
 OS Hepatitis B virus (subtype adw / strain Japan/DJMD233).
 OC Viruses; Retroviral viruses; Hepadnaviridae; Orthohepadnavirus.
 OX NCBI_TaxID=10413;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89010694; PubMed=3171552;
 RA Okamoto H., Tenda F., Sakugawa H., Sastroewiguo R.I., Imai M., Miyakawa Y., Mayumi M.,
 RT "Typing hepatitis B virus by homology in nucleotide sequence:
 RT comparison of surface antigen subtypes.";
 RL J. Gen. Virol. 69:2575-2583(1988).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: D00329; -; NOT_ANNOTATED_CDS.
 DR PIR: G28925; SAVLJ1.
 DR InterPro: IPR000349; Hepvlt_surfag.
 DR Pfam: PF00695; VMSA; 1.
 DR Antigen.
 KM Antigen.
 FT PROPEP 1 163
 FT CHAIN 164 389
 FT CARBOHYD 4 4
 FT CARBOHYD 112 112
 FT CARBOHYD 166 166
 FT CARBOHYD 309 309
 SQ SEQUENCE 389 AA; 42603 MW; 72F8D481924DC104 CRC64;

Query Match 88.9%; Score 1715; DB 1; Length 389;
 Best Local Similarity 89.5%; Pred. No. 1.5e-114;
 Matches 306; Conservative 12; Mismatches 24; Indels 0; Gaps 0;

QY 6 GAGAFGLGFTPPHGGILGMSPOAGILFTLPANPPASTNRSGROPTPLSPPLANTHPQ 65
 DB 48 GAGAFGGFTPPHGGILGMSPOAGILFTLPANPPASTNRSGROPTPLSPPLANTHPQ 107
 QY 66 AMONNSTTFHOTLADPRVRLYFPAGSSSGTVDVPTVSPISISIFSRIGDPALMMENT 125
 DB 108 AMONNSTTFHOTLADPRVRLYFPAGSSSGTVDVPTVSPISISISIRGDPALMMENT 167
 QY 126 TSGFLGPLLVLQAGFFLLTRILITIPQSLDSWMTSLNFIAGTTCVCLGONSOSPNSHPTS 185
 DB 168 ASGLGPLLVLQAGFFLLTRILITIPQSLDSWMTSLNFIAGTTCVCLGONSOSPNSHPTS 227
 QY 186 CPPTCGYRMWMLRRRIIFLLCLILFLVLDYQGLMPCVCLIPGSSSTTSGPCRTC 245
 DB 228 CPPTCGYRMWMLRRRIIFLLCLILFLVLDYQGLMPCVCLIPGSSSTTSGPCRTC 287
 QY 246 TTPAOGTSMYPSCCCTKPSDGNCTCIPIPSSWAFKFLMNASARFWSLTVFVQWFW 305
 DB 288 TTPAOGTSMYPSCCCTKPSDGNCTCIPIPSSWAFKFLMNASARFWSLTVFVQWFW 347
 QY 306 GLSPTVMTSLVIMMMWYMGPSLYSLSPPLPLPIFCLMVTYI 347
 DB 348 GLSPTVMTSLVIMMMWYMGPSLYSLSPPLPLPIFCLMVTYI 389

RESULT 9
 VMSA_HPBV2 STANDARD; PRT; 400 AA.

ID VMSA_HPBV2
 AC P03141;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Major surface antigen precursor.
 GN S.
 OS Hepatitis B virus (subtype adw2).
 OC Viruses; Retroviral viruses; Hepadnaviridae; Orthohepadnavirus.
 OX NCBI_TaxID=10408;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Valenzuela P., Quiroga M., Zaldivar J., Gray P., Rutter W.J.;
 RL (in) Field B.N., Jaenisch R., Fox C.F. (eds.);
 RL Animal virus genetics, pp.57-70, Academic Press, New York (1980).
 RN [2]
 RP SEQUENCE OF 175-400 FROM N.A.
 RX MEDLINE=79244739; PubMed=471053;
 RA Valenzuela P., Gray P., Quiroga M., Zaldivar J., Goodman H.M., Rutter W.J.;
 RT "Nucleotide sequence of the gene coding for the major protein of
 RT hepatitis B virus surface antigen.";
 RL Nature 280:815-819(1979).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X02763; CAA26539.1; -
 DR EMBL: J02205; AAA45524.1; -
 DR PIR: A03706; SAVLVD.
 DR InterPro: IPR000349; Hepvlt_surfag.
 DR Pfam: PF00695; VMSA; 1.
 DR Antigen.
 KM Antigen.
 FT PROPEP 1 174
 FT CHAIN 175 400
 FT CARBOHYD 15 15
 FT CARBOHYD 123 123
 FT CARBOHYD 177 177
 FT CARBOHYD 320 320
 SQ SEQUENCE 400 AA; 43704 MW; 57356B6293872BC5 CRC64;

Query Match 88.8%; Score 1713; DB 1; Length 400;
 Best Local Similarity 88.6%; Pred. No. 2.2e-114;
 Matches 303; Conservative 14; Mismatches 25; Indels 0; Gaps 0;

QY 6 GAGAGCGTTPPHGGGLDMSPOAGILETLTPANPPASTNROSGROPTPLSPPLRTHPQ 65
 DB 59 GAGAGCGTTPPHGGGLDMSPOAGILETLTPANPPASTNROSGROPTPLSPPLRTHPQ 118
 QY 66 AMQMNSTTFHQTLQDPRVAGLYFPAGSSSGCTVNPVTPSPISISFRIIDPALMENI 125
 DB 119 AMQMNSTTFHQTLQDPRVAGLYFPAGSSSGCTVNPVTPSPISISFRIIDPALMENI 178
 QY 126 TSGFLGFLVLAQGFLLRLITLTPQSLDSWMTSLNPLGCTTVCLGONSQSPNSHSPS 185
 DB 179 TSGFLGFLVLAQGFLLRLITLTPQSLDSWMTSLNPLGCTTVCLGONSQSPNSHSPS 238
 QY 186 CPPTCGYRMCRLRRIFLIFLLCLIFLVLLDYOGLPVCPLIPGSSSTTGPCRTC 245
 DB 239 CPPTCGYRMCRLRRIFLIFLLCLIFLVLLDYOGLPVCPLIPGSSSTTGPCRTC 298
 QY 246 TTPAGTSMYPSCCCTKPSDGNCTCIPISMAFGKFLMWMASARFSWLSLVPVQMFV 305
 DB 299 TTPAGTSMYPSCCCTKPSDGNCTCIPISMAFGKFLMWMASARFSWLSLVPVQMFV 358
 QY 306 GLSPTVMSLVIMMMWYGPSPYLSILSPFLPLIPFCLMWTYI 347
 DB 359 GLSPTVMSLVIMMMWYGPSPYLSILSPFLPLIPFCLMWTYI 400

RESULT 10

VMSA_HPBVO STANDARD; PRT; 389 AA.

ID VMSA_HPBVO
 AC P1739;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 01-AUG-1992 (Rel. 23, Last annotation update)
 DE Major surface antigen precursor.
 GN 5.
 OS Hepatitis B virus (subtype adw / strain Okinawa/p00W22).
 OC Viruses; Retrovird viruses; Hepadnaviridae; Orthohepadnavirus.
 OX NCBI_TaxID=10415;
 RN 11
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89010694; PubMed=3171552;
 RA Okamoto H., Tsuda F., Sakugawa H., Saetoeoewigao R.I., Imai M., Miyakawa Y., Mayumi M.;
 RT "Typing hepatitis B virus by homology in nucleotide sequence:
 RT comparison of surface antigen subtypes.";
 RL J. Gen. Virol. 69:2575-2583(1988)

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; D00330; -; NOT_ANNOTATED_CDS.

DR PIR; H28925; SAVLJ2.

DR InterPro; IPR000349; Hepvir_surfag.

DR Pfam; PF00695; VMSA; 1.

KW Antigen.

FT PROPEP 1 163

FT CHAIN 164 389

FT CARBOHYD 4 4

FT CARBOHYD 112 112

FT CARBOHYD 166 166

FT CARBOHYD 309 309

FT SEQUENCE 389 AA; 42408 MW; CECACODDA3D4A10 CRC64;

Matches 303; Conservative 15; Mismatches 24; Indels 0; Gaps 0;

QY 6 GAGAGCGTTPPHGGGLDMSPOAGILETLTPANPPASTNROSGROPTPLSPPLRTHPQ 65
 DB 48 GAGAGCGTTPPHGGGLDMSPOAGILETLTPANPPASTNROSGROPTPLSPPLRTHPQ 107
 QY 66 AMQMNSTTFHQTLQDPRVAGLYFPAGSSSGCTVNPVTPSPISISFRIIDPALMENI 125
 DB 108 AMQMNSTTFHQTLQDPRVAGLYFPAGSSSGCTVNPVTPSPISISFRIIDPALMENI 167
 QY 126 TSGFLGFLVLAQGFLLRLITLTPQSLDSWMTSLNPLGCTTVCLGONSQSPNSHSPS 185
 DB 168 TSGFLGFLVLAQGFLLRLITLTPQSLDSWMTSLNPLGCTTVCLGONSQSPNSHSPS 227
 QY 186 CPPTCGYRMCRLRRIFLIFLLCLIFLVLLDYOGLPVCPLIPGSSSTTGPCRTC 245
 DB 228 CPPTCGYRMCRLRRIFLIFLLCLIFLVLLDYOGLPVCPLIPGSSSTTGPCRTC 287
 QY 246 TTPAGTSMYPSCCCTKPSDGNCTCIPISMAFGKFLMWMASARFSWLSLVPVQMFV 305
 DB 288 TTPAGTSMYPSCCCTKPSDGNCTCIPISMAFGKFLMWMASARFSWLSLVPVQMFV 347
 QY 306 GLSPTVMSLVIMMMWYGPSPYLSILSPFLPLIPFCLMWTYI 347
 DB 348 GLSPTVMSLVIMMMWYGPSPYLSILSPFLPLIPFCLMWTYI 389

RESULT 11

VMSA_HPBVM STANDARD; PRT; 389 AA.

ID VMSA_HPBVM
 AC P03142;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Major surface antigen precursor.
 GN 5.
 OS Hepatitis B virus (subtype adw).
 OC Viruses; Retrovird viruses; Hepadnaviridae; Orthohepadnavirus.
 OX NCBI_TaxID=106821;
 RN 11
 RP SEQUENCE FROM N.A.
 RX MEDLINE=8316819; PubMed=6300776;
 RA Ono Y., Ono H., Sasada R., Igarashi K., Sugino Y., Nishio K.;
 RT "The complete nucleotide sequences of the cloned hepatitis B virus
 RT DNA; subtype adr and adw.";
 RL Nucleic Acids Res. 11:1747-1757(1983).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; V00866; CAA24233.1; ALT_INIT.

DR PIR; A93460; SAVLVE.

DR InterPro; IPR000349; Hepvir_surfag.

DR Pfam; PF00695; VMSA; 1.

KW Antigen.

FT PROPEP 1 163

FT CHAIN 164 389

FT CARBOHYD 4 4

FT CARBOHYD 112 112

FT CARBOHYD 166 166

FT CARBOHYD 309 309

FT SEQUENCE 389 AA; 42354 MW; FE9EP27BEFC58B5 CRC64;

Query Match 87.9%; Score 1695; DB 1; Length 389;

Best Local Similarity 87.7%; Pred. No. 4e-113;

Matches 300; Conservative 14; Mismatches 28; Indels 0; Gaps 0;


```

QY 6 GAGAFGLGFTPHHGLGMSPOAGILFTLPANPPASTNRQSGROPTPLSPPLRNTHPQ 65
DB 48 GVGAFFGLTTPHGGILGMSPOAGILFTVSTIPPPASTNRQSGROPTPLSPPLRNDHPQ 107
QY 66 AMQWSTTFHOTLQDPRVAGLYFPAGSSSGTVNVPVPTTSSISSIRICDPAIEMNI 125
DB 108 AMQWSTTFHOTLQDPRVAGLYFPAGSSSGTVNVPVPTTSSISSIRICDPAIEMNI 167
QY 126 TSGFLGPLLVLAQAGFFLRLITLITPOSIDSWMTSINFLAGTTVCIGQNSQSPHSPTS 185
DB 168 TSGFLGPLLVLAQAGFFLRLITLITPOSIDSWMTSINFLAGTTVCIGQNSQSPHSPTS 227
QY 186 CPPTCPGYRMMCLRRFIFLFLICLFLVLYLDYQMLPVCPILIGSSSTTSPTCTC 245
DB 228 CPPTCPGYRMMCLRRFIFLFLICLFLVLYLDYQMLPVCPILIGSSSTTSPTCTC 287
QY 246 TTPAOGTSMYPSCCCTKPSDGNCTCIPISSWAFKFLMENAARFWSLILVFPVQMFV 305
DB 288 TTPAOGTSMYPSCCCTKPSDGNCTCIPISSWAFKFLMENAARFWSLILVFPVQMFV 347
QY 306 GLSPTVWLSVIMMMWMPBSLYSLSPFLPLPIFFCLMAYI 347
DB 348 GLSPTVWLSVIMMMWMPBSLYSLSPFLPLPIFFCLMAYI 389

```

RESULT 12

```

VMSA_HPBVI STANDARD; PRT; 389 AA.
ID VMSA_HPBVI
AC P17357;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE Major surface antigen precursor.
GN S.
OS Hepatitis B virus (subtype adw / strain Indonesia/BIW420).
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
NCBI_TaxID=10412;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=89010694; PubMed=3171552;
RA Okamoto H., Tenda F., Sakugawa H., Saitosowejnjo R.I., Imai M.,
RA Miyakawa Y., Mayumi M.;
RT "Typing hepatitis B virus by homology in nucleotide sequence:
RT comparison of surface antigen subtypes."
RL J. Gen. Virol. 69:2575-2583(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, D00311; -; NOT_ANNOTATED_CDS.
DR PIR, I28925; SAVLJ3.
DR InterPro: IPR00349; Hepvir_surfag.
DR Pfam: PF00695; VMSA, 1.
KW Antigen.
FT PROPEP 1 163
FT CHAIN 164 389
FT CARBOHYD 4 4
FT CARBOHYD 112 112
FT CARBOHYD 166 166
FT CARBOHYD 309 309
SQ SEQUENCE 389 AA; 42653 MW; 08079BA34F3B90C0 CRC64;

```

```

Query Match 87.3%; Score 1684; DB 1; Length 389;
Best Local Similarity 88.0%; Pred. No. 2,4e-112;
Matches 301; Conservative 12; Mismatches 29; Indels 0; Gaps 0;

```

```

QY 6 GAGAFGLGFTPHHGLGMSPOAGILFTLPANPPASTNRQSGROPTPLSPPLRNTHPQ 65
DB 48 GVGAFFGLTTPHGGILGMSPOAGILFTVSTIPPPASTNRQSGROPTPLSPPLRNDHPQ 107
QY 66 AMQWSTTFHOTLQDPRVAGLYFPAGSSSGTVNVPVPTTSSISSIRICDPAIEMNI 125
DB 108 AMQWSTTFHOTLQDPRVAGLYFPAGSSSGTVNVPVPTTSSISSIRICDPAIEMNI 167
QY 126 TSGFLGPLLVLAQAGFFLRLITLITPOSIDSWMTSINFLAGTTVCIGQNSQSPHSPTS 185
DB 168 TSGFLGPLLVLAQAGFFLRLITLITPOSIDSWMTSINFLAGTTVCIGQNSQSPHSPTS 227
QY 186 CPPTCPGYRMMCLRRFIFLFLICLFLVLYLDYQMLPVCPILIGSSSTTSPTCTC 245
DB 228 CPPTCPGYRMMCLRRFIFLFLICLFLVLYLDYQMLPVCPILIGSSSTTSPTCTC 287
QY 246 TTPAOGTSMYPSCCCTKPSDGNCTCIPISSWAFKFLMENAARFWSLILVFPVQMFV 305
DB 288 TTPAOGTSMYPSCCCTKPSDGNCTCIPISSWAFKFLMENAARFWSLILVFPVQMFV 347
QY 306 GLSPTVWLSVIMMMWMPBSLYSLSPFLPLPIFFCLMAYI 347
DB 348 GLSPTVWLSVIMMMWMPBSLYSLSPFLPLPIFFCLMAYI 389

```

```

DB 48 GVGAFFGLTTPHGGILGMSPOAGILFTVSTIPPPASTNRQSGROPTPLSPPLRNTHPQ 107
QY 66 AMQWSTTFHOTLQDPRVAGLYFPAGSSSGTVNVPVPTTSSISSIRICDPAIEMNI 125
DB 108 AMQWSTTFHOTLQDPRVAGLYFPAGSSSGTVNVPVPTTSSISSIRICDPAIEMNI 167
QY 126 TSGFLGPLLVLAQAGFFLRLITLITPOSIDSWMTSINFLAGTTVCIGQNSQSPHSPTS 185
DB 168 TSGFLGPLLVLAQAGFFLRLITLITPOSIDSWMTSINFLAGTTVCIGQNSQSPHSPTS 227
QY 186 CPPTCPGYRMMCLRRFIFLFLICLFLVLYLDYQMLPVCPILIGSSSTTSPTCTC 245
DB 228 CPPTCPGYRMMCLRRFIFLFLICLFLVLYLDYQMLPVCPILIGSSSTTSPTCTC 287
QY 246 TTPAOGTSMYPSCCCTKPSDGNCTCIPISSWAFKFLMENAARFWSLILVFPVQMFV 305
DB 288 TTPAOGTSMYPSCCCTKPSDGNCTCIPISSWAFKFLMENAARFWSLILVFPVQMFV 347
QY 306 GLSPTVWLSVIMMMWMPBSLYSLSPFLPLPIFFCLMAYI 347
DB 348 GLSPTVWLSVIMMMWMPBSLYSLSPFLPLPIFFCLMAYI 389

```

RESULT 13

```

VMSA_HPBVI STANDARD; PRT; 400 AA.
ID VMSA_HPBVI
AC P17101;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE Major surface antigen precursor.
GN S.
OS Hepatitis B virus (subtype adw / strain 991).
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
NCBI_TaxID=10410;
RN (1)
RP SEQUENCE FROM N.A.
RX Koechel H.G., Schueler A., Lottmann S., Thomassen R.;
RA Submitted (FEB-1990) to the EMBL/GenBank/DBS databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, X51970; CAA36230.1; -
DR PIR, S10383; SAVLKS.
DR InterPro: IPR00349; Hepvir_surfag.
DR Pfam: PF00695; VMSA, 1.
KW Antigen.
FT PROPEP 1 174
FT CHAIN 175 400
FT CARBOHYD 15 15
FT CARBOHYD 123 123
FT CARBOHYD 177 177
FT CARBOHYD 320 320
SQ SEQUENCE 400 AA; 43695 MW; 76D98F748B17724 CRC64;

```

```

Query Match 87.1%; Score 1684; DB 1; Length 400;
Best Local Similarity 87.1%; Pred. No. 2,5e-112;
Matches 298; Conservative 16; Mismatches 28; Indels 0; Gaps 0;

```

```

QY 6 GAGAFGLGFTPHHGLGMSPOAGILFTLPANPPASTNRQSGROPTPLSPPLRNTHPQ 65
DB 59 GVGAFFGLTTPHGGILGMSPOAGILFTVSTIPPPASTNRQSGROPTPLSPPLRNDHPQ 118
QY 66 AMQWSTTFHOTLQDPRVAGLYFPAGSSSGTVNVPVPTTSSISSIRICDPAIEMNI 125
DB 119 AMQWSTTFHOTLQDPRVAGLYFPAGSSSGTVNVPVPTTSSISSIRICDPAIEMNI 178

```

```

OY 126 TSGFLGLPLVLAQGFLLTRILTIPOSIDSWMTSLNPLGTTVCLGONSOSPISNHSPTS 185
DB 179 TSGFLGLPLVLAQGFLLTRILTIPOSIDSWMTSLNPLGSPVCLGONSOSPISNHSPTS 238
OY 186 CPPTCGYRMWCLRRFIIIFLLILLCILFLLVLDYQGMPLVPCILIGSSSTSTGSPRTG 245
DB 239 CPPTCGYRMWCLRRFIIIFLLILLCILFLLVLDYQGMPLVPCILIGSSSTSTGSPRTG 298
OY 246 TTPAGCTMYSPSCCTKPSDGNCTCIPISSWAFGKFLMENASARFWSLILVPEVQNFV 305
DB 299 TTPAGCTMYSPSCCTKPSDGNCTCIPISSWAFGKFLMENASARFWSLILVPEVQNFV 358
OY 306 GLSPTVWLTVIMMMWYMGPSLYSLSPFLPLPIFFCLMWTI 347
DB 359 GLSPTVWLTVIMMMWYMGPSLYSLSPFLPLPIFFCLMWTI 400

RESULT 14
VMSA_HPBVT STANDARD; PRT; 400 AA.
AC 005486;
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Major surface antigen precursor.
GN S.
OS Hepatitis B virus (subtype adw4 / strain Brazil / isolate w4B).
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
NCBI_TaxID=45410;
RX MEDLINE=93346970; PubMed=8345355;
RA Naumann H., Schaefer S., Yoshida C.F.T., Gaspar A.M.C., Repp R.,
RA Gerlich W.H.;
RT "Identification of a new hepatitis B virus (HBV) genotype from Brazil
RT that expresses HBV surface antigen subtype adw4."
J. Gen. Virol. 74:1627-1632(1993).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; X69798; CAA49455.1; -
DR InterPro: IPR000349; Hepvir_surFag.
DR Pfam: PF00695; VMSA; 1.
KW Antigen.
FT PROPEP 175 400 MAJOR SURFACE ANTIGEN.
FT CHAIN 175 400
SQ SEQUENCE 400 AA; 43551 MW; 4A5A212E4B3E117 CRC64;

Query Match 83.7%; Score 1615; DB 1; Length 400;
Best Local Similarity 84.5%; Pred. No. 1,9e-107; Indels 0; Gaps 0;
Matches 289; Conservative 20; Mismatches 33;

OY 6 GAGAFGLGFTPPHGLGLMSPOAGILLETLPANPPASTNFSQSRQPTLSPPLRANTHPQ 65
DB 59 GAGGCGPFTPPHGLGLMSPOAGVLTLPADPPASTNRSRKRKPPVSPPLRDHPQ 118
OY 66 AMONNSTITFQTLDPFRVRLGYFPAGSSSGTVPNPVTPVPSISIFSRIGDPLAMENI 125
DB 119 AMONNSTITFQTLDPFRVRLGYFPAGSSSGTVPNPVTPVPSISIFSRIGDPLAMENI 178
OY 126 TSGFLGLPLVLAQGFLLTRILTIPOSIDSWMTSLNPLGTTVCLGONSOSPISNHSPTS 185
DB 179 TSGFLGLPLVLAQGFLLTRILTIPOSIDSWMTSLNPLGTCGCGONSOSPISNHSPTS 238
OY 186 CPPTCGYRMWCLRRFIIIFLLILLCILFLLVLDYQGMPLVPCILIGSSSTSTGSPRTG 245
DB 239 CPPTCGYRMWCLRRFIIIFLLILLCILFLLVLDYQGMPLVPCILIGSSSTSTGSPRTG 298

```

```

OY 246 TTPAGCTMYSPSCCTKPSDGNCTCIPISSWAFGKFLMENASARFWSLILVPEVQNFV 305
DB 299 TTPAGCTMYSPSCCTKPSDGNCTCIPISSWAFGKFLMENASARFWSLILVPEVQNFV 358
OY 306 GLSPTVWLTVIMMMWYMGPSLYSLSPFLPLPIFFCLMWTI 347
DB 359 GLSPTVWLTVIMMMWYMGPSLYSLSPFLPLPIFFCLMWTI 400

RESULT 15
VMSA_HPBVT STANDARD; PRT; 226 AA.
AC P31873;
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE Major surface antigen.
GN S.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
NCBI_TaxID=10407;
RX MEDLINE=88297159; PubMed=2841200;
RA Rivkina M.B., Luntin V.G., Mahov A.M., Tikhonchenko T.I., Kukain R.A.;
RT "Nucleotide sequence of integrated hepatitis B virus DNA and human
RT flanking regions in the genome of the PLC/PRF/5 cell line."
J. Gene 64:285-296(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; M21030; AAA45516.1; -
DR EMBL; X04820; CAA28506.1; -
DR PIR; J02933; SAVI4V.
DR InterPro: IPR000349; Hepvir_surFag.
DR Pfam: PF00695; VMSA; 1.
KW Antigen.
FT CARBOHYD 146 146 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 146 146 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 226 AA; 25476 MW; BFC4329CF1720600 CRC64;

Query Match 62.9%; Score 1214; DB 1; Length 226;
Best Local Similarity 93.4%; Pred. No. 2.5e-79; Indels 0; Gaps 0;
Matches 211; Conservative 9; Mismatches 6;

OY 122 MENITSGFLPLVLAQGFLLTRILTIPOSIDSWMTSLNPLGTTVCLGONSOSPISNHSPTS 181
DB 1 MENITSGFLPLVLAQGFLLTRILTIPOSIDSWMTSLNPLGSPVCLGONSOSPISNHSPTS 238
OY 182 SPTSCPTCGYRMWCLRRFIIIFLLILLCILFLLVLDYQGMPLVPCILIGSSSTSTGSP 241
DB 61 SPTSCPTCGYRMWCLRRFIIIFLLILLCILFLLVLDYQGMPLVPCILIGSSSTSTGSP 120
OY 242 CRTCTTTPAGCTMYSPSCCTKPSDGNCTCIPISSWAFGKFLMENASARFWSLILVPEV 301
DB 121 CRTCTTTPAGCTMYSPSCCTKPSDGNCTCIPISSWAFGKFLMENASARFWSLILVPEV 180
OY 302 QMFVGLSPTVWLTVIMMMWYMGPSLYSLSPFLPLPIFFCLMWTI 347
DB 181 QMFVGLSPTVWLTVIMMMWYMGPSLYSLSPFLPLPIFFCLMWTI 226

Search completed: February 3, 2003, 09:24:18
Job time : 14.6951 secs

```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 09:21:57; Search time 22.9721 Seconds

(without alignments)
1452.136 Million cell updates/sec

Title: US-09-890-752A-1

Sequence: 1 MRRGDGAGFGIGFTRPHGG.....SILSPPLPLPIPTCLWYI 347

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1901	98.5	389	1	SAVLBH
2	1883	97.6	389	2	S20745
3	1879	97.4	389	2	S32202
4	1875	97.2	389	2	S47407
5	1868	96.8	389	1	SAVLAH
6	1868	96.8	389	1	SAVLAI
7	1861	96.5	389	2	S20753
8	1851	96.0	389	1	SAVLAI
9	1803	93.5	389	2	S41871
10	1799	93.3	378	2	S41869
11	1797	93.2	389	2	S20749
12	1771	91.8	400	1	SAVLAI
13	1755	91.0	389	1	SAVLAI
14	1745	90.5	400	2	S35528
15	1734	89.9	378	2	S41870
16	1723	89.3	389	1	S67506
17	1715	88.9	389	1	SAVLAI
18	1713	88.8	400	1	SAVLAI
19	1713	88.8	400	1	SAVLAI
20	1711	88.7	445	2	S43492
21	1704	88.3	389	1	SAVLAI
22	1699	88.1	389	2	S36542
23	1695	87.9	389	1	SAVLAI
24	1684	87.3	389	1	SAVLAI
25	1684	87.3	400	1	SAVLAI
26	1641	85.1	384	2	T13469
27	1615	83.7	400	2	T13474
28	1593	82.6	382	2	T13474
29	1324	68.6	261	2	J02226

30	1269	65.8	226	2	J02077	surface antigen -
31	1269	65.8	226	2	J02076	surface antigen -
32	1267	65.7	226	1	J01571	major surface anti
33	1265	65.6	226	2	J02075	surface antigen -
34	1264	65.5	226	2	J02078	surface antigen -
35	1264	65.5	226	2	J02078	surface antigen -
36	1263	65.5	226	2	J02069	surface antigen -
37	1260	65.3	226	2	J02073	surface antigen -
38	1259	65.3	226	2	J02067	surface antigen -
39	1257	65.2	226	2	J02063	surface antigen -
40	1256	65.1	226	2	J02063	surface antigen -
41	1253	65.0	226	1	J01572	major surface anti
42	1253	65.0	226	2	J02065	surface antigen -
43	1250	64.8	226	2	J02070	surface antigen -
44	1250	64.8	226	2	J02065	surface antigen -
45	1249	64.7	226	2	J02081	surface antigen -

ALIGNMENTS

RESULT 1

SAVLBH large surface antigen - hepatitis B virus (subtype ayw, strain PHB320)

M/Contains: major surface antigen; middle surface antigen
C/Species: hepatitis B virus, HBV
A/Note: host Homo sapiens (man)

C/Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 28-Jul-2000
C/Accession: A03704; P00585
R/Bibiko, V.; Peshko, P.; Drejlina, D.; Pumpen, P.; Gren, B.

FEBS Lett. 185, 208-212, 1985
A/Title: Subtype ayw variant of hepatitis B virus: DNA primary structure analysis.

A/Reference number: A05237; MUID:85204397; PMID:3996597

A/Accession: A03704

A/Molecule type: DNA

A/Residues: 1389 <BIC>
A/Cross-references: GB:X02496; MID:962280; PIDN:CAB41701.1; PID:94704321

A/Norder, H.; Courouce, A.M.; Magnius, L.O.
J. Gen. Virol. 73, 3141-3145, 1992
A/Title: Molecular basis of hepatitis B virus serotype variations within the four major

A/Reference number: P00453; MUID:93107848; PMID:1469353

A/Accession: P00585
A/Molecule type: DNA
A/Residues: 264-343 <NOR>
A/Experimental source: subtype ayw2, Tav

C/Genetics:
A/Gene: pre-S1/pre-S2/S
C/Superfamily: hepatitis B virus surface antigen

C/Keywords: glycoprotein; surface antigen
F/109-389/Product: middle surface antigen (gene pre-S2/S) #status predicted <MSA>
F/164-389/Product: major surface antigen (gene S) #status predicted <MSA>
F/4,112,166/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 98.5%; Score 1901; DB 1; Length 389;
Best Local Similarity 100.0%; Pred. No. 3.3e-18;

Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	6	GAGAFGLGFTPRPGGLGSPGPOGLTETLPANPPASTNRSGOPTPLSPPLKNTHQ	65
DB	48	GAGAFGLGFTPRPGGLGSPGPOGLTETLPANPPASTNRSGOPTPLSPPLKNTHQ	107
QY	66	ANQWNTTHTQLODRVAGLYFPAGSSSGTVNPVTSSISIFSRIGDPALMNTI	125
DB	108	ANQWNTTHTQLODRVAGLYFPAGSSSGTVNPVTSSISIFSRIGDPALMNTI	167
QY	126	TSGFGLPLVLVLAGFPLRLITLITPOSLSWTSLNPLGCTTVCAGQNSQSPFNSHSPTS	185
DB	168	TSGFGLPLVLVLAGFPLRLITLITPOSLSWTSLNPLGCTTVCAGQNSQSPFNSHSPTS	227
QY	186	CPPTCGYRWMLRRFIFLFLITLCLIFLVLVDYQGLPVCPLIPSSSTTSQPCRTC	245
DB	228	CPPTCGYRWMLRRFIFLFLITLCLIFLVLVDYQGLPVCPLIPSSSTTSQPCRTC	287

OY 246 TTPAGTSMYPSCCCTKPSDGNCTCIPSPSWAGKFLMENASARFSGMLSLVFPVQMFV 305
 DB 288 TTPAGTSMYPSCCCTKPSDGNCTCIPSPSWAGKFLMENASARFSGMLSLVFPVQMFV 347
 OY 306 GLSPPTVMSLVIMMMWYMGPSLYSLISPLPLPLPIFFCLMAYI 347
 DB 348 GLSPPTVMSLVIMMMWYMGPSLYSLISPLPLPLPIFFCLMAYI 389

RESULT 2

520745
 surface antigen - hepatitis B virus (subtype ayw, patient C)
 N:Alternate names: HBe antigen
 N:Contains: surface antigen pre-S1 (large envelope protein); surface antigen pre-S2 (mid
 C:Species: hepatitis B virus, HBV
 A:Variety: subtype ayw, patient C
 C:Date: 20-Feb-1995 #sequence_revision 01-Nov-1996 #text_change 26-Aug-1999
 C:Accession: S20745
 R:Lat, M.B.; Mazzoli, A.P.; Balestrieri, A.; Mellis, A.; Porru, A.
 A:Description: Sequence analysis of HBV genomes isolated from patients with HBsAg negati
 A:Reference number: S20745
 A:Accession: S20745
 A:Molecule type: DNA
 A:Residues: 1-389 <LAT>
 A:Cross-references: EMBL:X65257; NID:G59429; PIDN:CAA6349.1; PID:G59430
 A:Experimental source: subtype ayw, patient C
 C:Genetics:
 A:Gene: S
 C:Superfamily: hepatitis B virus surface antigen
 C:Keywords: surface antigen
 F:1-389/Product: surface antigen pre-S1 (large envelope protein) #status predicted <PS1>
 F:1-108/Domains: pre-S1 domain #status predicted <PRE1>
 F:109-389/Product: surface antigen pre-S2 (middle envelope protein) #status predicted <PS2>
 F:109-163/Domains: pre-S2 domain #status predicted <PRE2>
 F:164-389/Product: surface antigen S (small envelope protein) #status predicted <PSD>

Query Match 97.4%; Score 1883; DB 2; Length 389;

Best Local Similarity 98.8%; Pred. No. 6.3e-127;

Matches 338; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 6 GAGAFGLGFTPPHGGILGMSPOAGIILTLPANPPASTNRSGRQPTPLSPRLRTHQ 65
 DB 48 GAGAFGLGFTPPHGGILGMSPOAGIILTLPANPPASTNRSGRQPTPLSPRLRTHQ 107
 OY 66 AMONSTTFHOTLDDPRVRLGLYPAGSSSGTVPPTVSPISIFSRIGDPAIMNENI 125
 DB 108 AMONSTTFHOTLDDPRVRLGLYPAGSSSGTVPPTVSPISIFSRIGDPAIMNENI 167
 OY 126 TSGFLGLVLYOAGFLLTRILITPQSLDSWMTSLNPLGTTVCLGNSQSPSNNHPTS 185
 DB 168 TSGFLGLVLYOAGFLLTRILITPQSLDSWMTSLNPLGTTVCLGNSQSPSNNHPTS 227
 OY 186 CPTCPGRRMCLRRFIIFLLCLIFLLVLDYOGMLPVCPLIPGSSSTTSGPCRTC 245
 DB 228 CPTCPGRRMCLRRFIIFLLCLIFLLVLDYOGMLPVCPLIPGSSSTTSGPCRTC 287
 OY 246 TTPAGTSMYPSCCCTKPSDGNCTCIPSPSWAGKFLMENASARFSGMLSLVFPVQMFV 305
 DB 288 TTPAGTSMYPSCCCTKPSDGNCTCIPSPSWAGKFLMENASARFSGMLSLVFPVQMFV 347
 OY 306 GLSPPTVMSLVIMMMWYMGPSLYSLISPLPLPLPIFFCLMAYI 347
 DB 348 GLSPPTVMSLVIMMMWYMGPSLYSLISPLPLPLPIFFCLMAYI 389

RESULT 3

S32202
 large surface antigen - hepatitis B virus (subtype ayw, isolate patient C1005)
 N:Contains: major surface antigen; middle surface antigen
 C:Species: hepatitis B virus, HBV
 A:Variety: subtype ayw, isolate patient C1005
 C:Date: 20-Feb-1995 #sequence_revision 06-Dec-1996 #text_change 06-Dec-1996

C:Accession: S32202

R:Preiser-Adams, S.; Schlayer, M.J.; Peters, T.; Hettler, F.; Gerok, W.; Rasenack, J.

A:Submitted to the EMBL Data Library, March 1993

A:Description: Identification and sequence analysis of hepatitis B virus DNA in immunolog

A:Reference number: S32202

A:Accession: S32202

A:Molecule type: DNA

A:Residues: 1-389 <PRE>

A:Cross-references: EMBL:X72702

A:Experimental source: subtype ayw, isolate patient C1005

C:Genetics:

A:Gene: pre-S1/pre-S2/S

C:Superfamily: hepatitis B virus surface antigen

C:Keywords: surface antigen

F:1-389/Product: large surface antigen (gene pre-S1/pre-S2/S) #status predicted <DSL>

F:109-389/Product: middle surface antigen (gene pre-S2/S) #status predicted <DSA>

F:164-389/Product: major surface antigen (gene S) #status predicted <MSA>

Query Match 97.4%; Score 1879; DB 2; Length 389;

Best Local Similarity 98.5%; Pred. No. 1.2e-126;

Matches 337; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 6 GAGAFGLGFTPPHGGILGMSPOAGIILTLPANPPASTNRSGRQPTPLSPRLRTHQ 65
 DB 48 GAGAFGLGFTPPHGGILGMSPOAGIILTLPANPPASTNRSGRQPTPLSPRLRTHQ 107
 OY 66 AMONSTTFHOTLDDPRVRLGLYPAGSSSGTVPPTVSPISIFSRIGDPAIMNENI 125
 DB 108 AMONSTTFHOTLDDPRVRLGLYPAGSSSGTVPPTVSPISIFSRIGDPAIMNENI 167
 OY 126 TSGFLGLVLYOAGFLLTRILITPQSLDSWMTSLNPLGTTVCLGNSQSPSNNHPTS 185
 DB 168 TSGFLGLVLYOAGFLLTRILITPQSLDSWMTSLNPLGTTVCLGNSQSPSNNHPTS 227
 OY 186 CPTCPGRRMCLRRFIIFLLCLIFLLVLDYOGMLPVCPLIPGSSSTTSGPCRTC 245
 DB 228 CPTCPGRRMCLRRFIIFLLCLIFLLVLDYOGMLPVCPLIPGSSSTTSGPCRTC 287

OY 246 TTPAGTSMYPSCCCTKPSDGNCTCIPSPSWAGKFLMENASARFSGMLSLVFPVQMFV 305

DB 288 TTPAGTSMYPSCCCTKPSDGNCTCIPSPSWAGKFLMENASARFSGMLSLVFPVQMFV 347

OY 306 GLSPPTVMSLVIMMMWYMGPSLYSLISPLPLPLPIFFCLMAYI 347

DB 348 GLSPPTVMSLVIMMMWYMGPSLYSLISPLPLPLPIFFCLMAYI 389

RESULT 4

S47407

surface antigen - hepatitis B virus (subtype ayw4)

N:Alternate names: envelope protein, HBe antigen

N:Contains: surface antigen pre-S1 (large envelope protein); surface antigen pre-S2 (mid

C:Species: hepatitis B virus, HBV

A:Variety: subtype ayw4

C:Date: 23-Nov-1994 #sequence_revision 12-May-1995 #text_change 26-Aug-1999

C:Accession: S47407

A:Description: Molecular cloning and sequencing of two complete genomes of polish isolate

A:Reference number: S47404

A:Accession: S47407

A:Molecule type: DNA

A:Residues: 1-389 <PLU>

A:Cross-references: EMBL:235716; NID:G527435; PIDN:CAA64768.1; PID:G527439

A:Experimental source: subtype ayw4

R:Rorder, H.; Hammas, B.; Lee, S.D.; Bille, K.; Courouce, A.M.; Mushahwar, I.K.; Magnus,

U. Gen. Virol. 74, 1341-1348, 1993

A:Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin

A:Reference number: JQ2044; MUID:93329382; PMID:8336122

A:Accession: JQ2071

A:Molecule type: DNA

A:Residues: 164-389 <NOR>

A:Experimental source: subtype ayw3, strain 8950/90
 C:Genetics:
 A:Gene: S
 A:Introns: 111/3
 C:Superfamily: hepatitis B virus surface antigen
 C:Keywords: glycoprotein; surface antigen
 F:109-389/Product: surface antigen
 F:1-108/Domain: pre-S1 domain #status predicted <PS1>
 F:109-389/Product: surface antigen pre-S2 (middle envelope protein) #status predicted <PS2>
 F:109-163/Domain: pre-S2 domain #status predicted <PS2>
 F:164-389/Product: surface antigen S (small envelope protein) #status predicted <SAG>

Query Match
 Best Local Similarity 97.2%; Score 1875; DB 2; Length 389;
 Matches 336; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 6 GAGAGLGLFTPPHGLGMSPOAGILETLPANPPASTNRSGROPTPLSPRLNTHPQ 65
 DB 48 GAGAGLGLFTPPHGLGMSPOAGILETLPANPPASTNRSGROPTPLSPRLNTHPQ 107
 QY 66 AMONSTTHQTLQDPVRGLYFPAGSSSGTVNPVTASPISSIFSRIGDPALNMEI 125
 DB 108 AMONSTTHQTLQDPVRGLYFPAGSSSGTVNPVTASPISSIFSRIGDPALNMEI 167
 QY 126 TSGFLGPLVLAQAFPLRLITLTPQSLDSWTSINLVGTTVCLGONSQSPSNHSPTS 185
 DB 168 TSGFLGPLVLAQAFPLRLITLTPQSLDSWTSINLVGTTVCLGONSQSPSNHSPTS 227
 QY 186 CPPTCPGYRMWMLARFIIFLLCLIFLLVLDYQGLPVCEPLIPGSSSTTSGPCRTC 245
 DB 228 CPPTCPGYRMWMLARFIIFLLCLIFLLVLDYQGLPVCEPLIPGSSSTTSGPCRTC 287
 QY 246 TTPAGTSMYPPCCCTKPSDGNCTCIPPSMAFGKFLMWSARFSLVFPVQMFV 305
 DB 288 TTPAGTSMYPPCCCTKPSDGNCTCIPPSMAFGKFLMWSARFSLVFPVQMFV 347
 QY 306 GLSPTVWLSVIMMMWYMGPSLYSLSPFLPLPIFCLMWYI 347
 DB 348 GLSPTVWLSVIMMMWYMGPSLYSLSPFLPLPIFCLMWYI 389

RESULT 5

SAVIAH

large surface antigen - hepatitis B virus (subtype ayw)
 N:Contains: major surface antigen; middle surface antigen
 C:Species: hepatitis B virus, HBV
 C>Date: 18-Dec-1981 #sequence, revision 18-Dec-1981 #text_change 07-May-1999
 C:Accession: A03703; JQ2064; PQ0591
 R:Galibert, F.; Mandart, E.; Fitoussi, F.; Tiollais, P.; Charnay, P.
 Nature 281, 646-650, 1979
 A:Title: Nucleotide sequence of the hepatitis B virus genome (subtype ayw) in E. coli.
 A:Reference number: A93214; MUID:81012091; PMID:399327
 A:Accession: A03703
 A:Molecule type: DNA
 A:Residues: 1-389 <GAL>
 A:Cross-references: GB:J02203
 R:Nordest, H.; Hammas, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnus, J. Gen. Virol. 74, 1341-1348, 1993
 A:Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
 A:Reference number: JQ2044; MUID:93329382; PMID:8336122
 A:Molecule type: DNA
 A:Residues: 164-389 <NOR>
 A:Experimental source: genogroup D, subtype ayw3
 R:Nordest, H.; Courouce, A.M.; Magnus, U.O.
 J. Gen. Virol. 73, 3141-3145, 1992
 A:Title: Molecular basis of hepatitis B virus serotype variations within the four major
 A:Reference number: PQ0453; MUID:93107848; PMID:1469353
 A:Molecule type: DNA
 A:Residues: 264-343 <NO2>
 A:Experimental source: subtype ayw3, K11
 C:Genetics:

A:Gene: pre-S1/pre-S2/S
 C:Superfamily: hepatitis B virus surface antigen
 C:Keywords: glycoprotein; surface antigen (gene pre-S2/S) #status predicted <PSA>
 F:109-389/Product: middle surface antigen (gene S) #status predicted <MSA>
 F:164-389/Product: major surface antigen (gene S) #status predicted <MSA>
 F:4,112,166/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match
 Best Local Similarity 96.8%; Score 1868; DB 1; Length 389;
 Matches 336; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 6 GAGAGLGLFTPPHGLGMSPOAGILETLPANPPASTNRSGROPTPLSPRLNTHPQ 65
 DB 48 GAGAGLGLFTPPHGLGMSPOAGILETLPANPPASTNRSGROPTPLSPRLNTHPQ 107
 QY 66 AMONSTTHQTLQDPVRGLYFPAGSSSGTVNPVTASPISSIFSRIGDPALNMEI 125
 DB 108 AMONSTTHQTLQDPVRGLYFPAGSSSGTVNPVTASPISSIFSRIGDPALNMEI 167
 QY 126 TSGFLGPLVLAQAFPLRLITLTPQSLDSWTSINLVGTTVCLGONSQSPSNHSPTS 185
 DB 168 TSGFLGPLVLAQAFPLRLITLTPQSLDSWTSINLVGTTVCLGONSQSPSNHSPTS 227
 QY 186 CPPTCPGYRMWMLARFIIFLLCLIFLLVLDYQGLPVCEPLIPGSSSTTSGPCRTC 245
 DB 228 CPPTCPGYRMWMLARFIIFLLCLIFLLVLDYQGLPVCEPLIPGSSSTTSGPCRTC 287
 QY 246 TTPAGTSMYPPCCCTKPSDGNCTCIPPSMAFGKFLMWSARFSLVFPVQMFV 305
 DB 288 TTPAGTSMYPPCCCTKPSDGNCTCIPPSMAFGKFLMWSARFSLVFPVQMFV 347
 QY 306 GLSPTVWLSVIMMMWYMGPSLYSLSPFLPLPIFCLMWYI 347
 DB 348 GLSPTVWLSVIMMMWYMGPSLYSLSPFLPLPIFCLMWYI 389

RESULT 6

SAVIAH

large surface antigen - hepatitis B virus (strain alpha1)
 N:Contains: major surface antigen; middle surface antigen
 C:Species: hepatitis B virus, HBV
 C>Date: 30-Sep-1991 #sequence, revision 30-Sep-1991 #text_change 16-Jul-1999
 C:Accession: B34773; JQ2082
 R:Tong, S.; Li, J.; Vitvitski, L.; Trepo, C.
 Virology 176, 596-603, 1990
 A:Title: Active hepatitis B virus replication in the presence of anti-HBe is associated v
 A:Reference number: A34773; MUID:9026476; PMID:2345966
 A:Accession: B34773
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-389 <TON>
 A:Cross-references: EMBL:M32138; NID:G329667; PIDN:AAA45502.1; PID:G329669
 J. Gen. Virol. 74, 1341-1348, 1993
 A:Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
 A:Reference number: JQ2044; MUID:93329382; PMID:8336122
 A:Molecule type: DNA
 A:Residues: 164-389 <NOR>
 A:Experimental source: genogroup D, subtype ayw2, strain HBV-alpha
 C:Genetics:
 A:Gene: pre-S1/pre-S2/S
 A:Introns: 111/3
 C:Superfamily: hepatitis B virus surface antigen
 C:Keywords: glycoprotein; surface antigen
 F:109-389/Product: middle surface antigen (gene pre-S2/S) #status predicted <PSA>
 F:164-389/Product: major surface antigen (gene S) #status predicted <MSA>
 F:4,112,166/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match
 Best Local Similarity 96.8%; Score 1868; DB 1; Length 389;
 Matches 334; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

```

Qy 6 GAGAFGLGFTPPHGGGLGMSPOAGIILETLPANPPASTNRSGRQPTPLSPPLRNTHPQ 65
Db 48 GAGAFGLGFTPPHGGGLGMSPOAGIILETLPANPPASTNRSGRQPTPLSPPLRNTHPQ 107
Qy 66 AMONNSTTFHQTLODPRVRGLYPPAGSSGCTVNPVTTSPISISIRKIGDPLANNENI 125
Db 108 AMONNSTTFHQTLODPRVRGLYPPAGSSGCTVNPVTTSPISISIRKIGDPLANNENI 167
Qy 126 TSGFGLPLVVOAGFPLTRILITIPQSLDSWMTSLNFIAGTTVCLGQNSQSPISNHSPTS 185
Db 168 TSGFGLPLVVOAGFPLTRILITIPQSLDSWMTSLNFIAGTTVCLGQNSQSPISNHSPTS 227
Qy 186 CPPTCPGYRMWCLRRFIIIFLLILLCILFLLVLDYQMLPVCPLIPGSSSTTSGPCRTC 245
Db 228 CPPTCPGYRMWCLRRFIIIFLLILLCILFLLVLDYQMLPVCPLIPGSSSTTSGPCRTC 287
Qy 246 TTPAGGTSMYPSGCCCTKXSDGNCCTCIPIPSSNAFGKFLMEWASARFWSLSLVFVQMFV 305
Db 288 TTPAGGTSMYPSGCCCTKXSDGNCCTCIPIPSSNAFGKFLMEWASARFWSLSLVFVQMFV 347
Qy 306 GLSPTVWLSVIMMMWYWGPSLSYLSLSPFLPLPIFFCLMAYI 347
Db 348 GLSPTVWLSVIMMMWYWGPSLSYLSLSPFLPLPIFFCLMAYI 389

```

RESULT 7

```

S20753
surface antigen - hepatitis B virus (subtype ayw, patient E)
N:Alternate names: envelope protein, HBe antigen
N:Contains: surface antigen pre-S1 (large envelope protein); surface antigen pre-S2 (mid
C:Species: hepatitis B virus, HBV
A:Variety: subtype ayw, patient E
C:Date: 20-Feb-1995 #sequence_revision 01-Nov-1996 #text_change 26-Aug-1999
C:Accession: S20753
A:Reference number: S20745
A:Molecule type: DNA
A:Residues: 1-389 <LAI>
A:Cross-references: EMBL:X65259; NID:G59439; PIDN:CAA46357.1; PID:G59440
A:Experimental source: subtype ayw, patient E
C:Genetics:
A:Gene: S
C:Superfamily: hepatitis B virus surface antigen
C:Keywords: surface antigen
F:1-108/Domain: pre-S1 domain #status predicted <PS1>
F:1-108/Domain: pre-S1 domain #status predicted <PS1>
F:109-389/Product: surface antigen pre-S2 (middle envelope protein) #status predicted <F
F:109-389/Domain: pre-S2 domain #status predicted <PS2>
F:164-389/Product: surface antigen S (small envelope protein) #status predicted <ESD>

```

Query Match 96.5%; Score 1861; DB 2; Length 389;
Best Local Similarity 97.7%; Pred. No. 2, 3e-125;
Matches 334; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```

Qy 6 GAGAFGLGFTPPHGGGLGMSPOAGIILETLPANPPASTNRSGRQPTPLSPPLRNTHPQ 65
Db 48 GAGAFGLGFTPPHGGGLGMSPOAGIILETLPANPPASTNRSGRQPTPLSPPLRNTHPQ 107
Qy 66 AMONNSTTFHQTLODPRVRGLYPPAGSSGCTVNPVTTSPISISIRKIGDPLANNENI 125
Db 108 AMONNSTTFHQTLODPRVRGLYPPAGSSGCTVNPVTTSPISISIRKIGDPLANNENI 167
Qy 126 TSGFGLPLVVOAGFPLTRILITIPQSLDSWMTSLNFIAGTTVCLGQNSQSPISNHSPTS 185
Db 168 TSGFGLPLVVOAGFPLTRILITIPQSLDSWMTSLNFIAGTTVCLGQNSQSPISNHSPTS 227
Qy 186 CPPTCPGYRMWCLRRFIIIFLLILLCILFLLVLDYQMLPVCPLIPGSSSTTSGPCRTC 245
Db 228 CPPTCPGYRMWCLRRFIIIFLLILLCILFLLVLDYQMLPVCPLIPGSSSTTSGPCRTC 287

```

```

Qy 246 TTPAGGTSMYPSGCCCTKXSDGNCCTCIPIPSSNAFGKFLMEWASARFWSLSLVFVQMFV 305
Db 288 TTPAGGTSMYPSGCCCTKXSDGNCCTCIPIPSSNAFGKFLMEWASARFWSLSLVFVQMFV 347
Qy 306 GLSPTVWLSVIMMMWYWGPSLSYLSLSPFLPLPIFFCLMAYI 347
Db 348 GLSPTVWLSVIMMMWYWGPSLSYLSLSPFLPLPIFFCLMAYI 389

```

RESULT 8

```

SAVLAD
large surface antigen - hepatitis B virus (subtype adyw)
N:Contains: major surface antigen; middle surface antigen
C:Species: hepatitis B virus, HBV
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 15-Nov-1996
C:Accession: A93217; A03703
R:Parsek, M.; Goto, T.; Gilbert, W.; Zink, B.; Schaller, H.; Mackay, P.; Leadbetter, G.;
Nature 282, 575-579, 1979
A:Title: Hepatitis B virus genes and their expression in E. coli.
A:Reference number: A93217; MUID:81012115; PMID:399329
A:Accession: A93217
A:Molecule type: DNA
A:Residues: 1-389 <PAS>
A:Cross-references: GB:J02202
C:Genetics:
A:Gene: pre-S1/pre-S2/S
C:Superfamily: hepatitis B virus surface antigen
C:Keywords: glycoprotein; surface antigen (gene pre-S2/S) #status predicted <DNA>
F:109-389/Product: middle surface antigen (gene S) #status predicted <MSA>
F:164-389/Product: major surface antigen (gene S) #status predicted <MSA>
F:4,112,166/Binding site: carbohydrate (Asn) (covalent) #status predicted

```

Query Match 96.4%; Score 1851; DB 1; Length 389;

Best Local Similarity 97.1%; Pred. No. 1, 2e-124;
Matches 332; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

```

Qy 6 GAGAFGLGFTPPHGGGLGMSPOAGIILETLPANPPASTNRSGRQPTPLSPPLRNTHPQ 65
Db 48 GAGAFGLGFTPPHGGGLGMSPOAGIILETLPANPPASTNRSGRQPTPLSPPLRNTHPQ 107
Qy 66 AMONNSTTFHQTLODPRVRGLYPPAGSSGCTVNPVTTSPISISIRKIGDPLANNENI 125
Db 108 AMONNSTTFHQTLODPRVRGLYPPAGSSGCTVNPVTTSPISISIRKIGDPLANNENI 167
Qy 126 TSGFGLPLVVOAGFPLTRILITIPQSLDSWMTSLNFIAGTTVCLGQNSQSPISNHSPTS 185
Db 168 TSGFGLPLVVOAGFPLTRILITIPQSLDSWMTSLNFIAGTTVCLGQNSQSPISNHSPTS 227
Qy 186 CPPTCPGYRMWCLRRFIIIFLLILLCILFLLVLDYQMLPVCPLIPGSSSTTSGPCRTC 245
Db 228 CPPTCPGYRMWCLRRFIIIFLLILLCILFLLVLDYQMLPVCPLIPGSSSTTSGPCRTC 287
Qy 246 TTPAGGTSMYPSGCCCTKXSDGNCCTCIPIPSSNAFGKFLMEWASARFWSLSLVFVQMFV 305
Db 288 TTPAGGTSMYPSGCCCTKXSDGNCCTCIPIPSSNAFGKFLMEWASARFWSLSLVFVQMFV 347
Qy 306 GLSPTVWLSVIMMMWYWGPSLSYLSLSPFLPLPIFFCLMAYI 347
Db 348 GLSPTVWLSVIMMMWYWGPSLSYLSLSPFLPLPIFFCLMAYI 389

```

RESULT 9

```

S41871
surface antigen - hepatitis B virus (subtype ayw, patient P)
N:Alternate names: envelope protein, HBe antigen
N:Contains: surface antigen pre-S1 (large envelope protein); surface antigen pre-S2 (mid
C:Species: hepatitis B virus, HBV
A:Variety: subtype ayw, patient P
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 26-Aug-1999
C:Accession: S41871
R:Lat, M.E.; Mazzoleni, A.P.; Balestrieri, A.
Submitted to the EMBL Data Library, January 1994
A:Description: Sequence analysis of HBV genomes isolated from patients with HBeAg chroni..

```

A:Reference number: S41869
 A:Accession: S41871
 A:Molecule type: DNA
 A:Residues: 1389 <LAI>
 A:Cross-references: EMBL:X77310; NID:G452615; PIDN:CAA54516.1; PID:G452612
 A:Experimental source: subtype ayw, patient P
 C:Genetics:
 A:Gene: S
 C:Superfamily: hepatitis B virus surface antigen
 C:Keywords: surface antigen
 F:1-389/Product: surface antigen pre-S1 (large envelope protein) #status predicted <PS1>
 F:1-108/Domain: pre-S1 domain #status predicted <PRB1>
 F:109-378/Product: surface antigen pre-S2 (middle envelope protein) #status predicted <PS2>
 F:109-163/Domain: pre-S2 domain #status predicted <PRB2>
 F:164-389/Product: surface antigen S (small envelope protein) #status predicted <PSD>

Query Match 93.5%; Score 1803; DB 2; Length 389;
 Best Local Similarity 95.3%; Pred. No. 3.1e-121;
 Matches 326; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 6 GAGAFGLGFTPPHGGILGWSPOAGILLETLPANPPASTNRSGROPTPLSPPLANTHPQ 65
 DB 48 GAGAFGLGFTPPHGGILGWSPOAGILLETLPANPPASTNRSGROPTPLSPPLANTHPQ 107
 QY 66 AMQNMSTTFHOTLQDPRVRGLYFPAGSSSGTVNPVTVPISISIFSRIGDPALNMENI 125
 DB 108 AMQNMSTTFHOTLQDPRVRGLYFPAGSSSGTVNPVTVPISISIFSRIGDPALNMENI 167
 QY 126 TSGFGLPLVLOAGFFLLTRILITPQSLDSWMTSLNPLGCTTVCLGQNSQSPSNNHSPTS 185
 DB 168 TSGFGLPLVLOAGFFLLTRILITPQSLDSWMTSLNPLGCTTVCLGQNSQSPSNNHSPTS 227
 QY 186 CPPTCGYRMWCLRRFFILFLILLCIFLVLVDYQMLPVCPLIPGSSSTTSGPCRTC 245
 DB 228 CPPTCGYRMWCLRRFFILFLILLCIFLVLVDYQMLPVCPLIPGSSSTTSGPCRTC 287
 QY 246 TTPAGSTMYPSGCCCTKPSDNCCTCIPSSWAFGKFLMWSARFWSLSLVFVQMFV 305
 DB 288 TTPAGSTMYPSGCCCTKPSDNCCTCIPSSWAFGKFLMWSARFWSLSLVFVQMFV 347
 QY 306 GLSPTVWLSVIMMMWYMGPSLYSLISPLPLIPFCCLMVTI 347
 DB 348 GLSPTVWLSVIMMMWYMGPSLYSLISPLPLIPFCCLMVTI 389

RESULT 10
 S41869
 surface antigen - hepatitis B virus (subtype ayw, patient A) (fragment)
 N:Alternate names: envelope protein; HBs antigen
 N:Contains: surface antigen pre-S1 (large envelope protein); surface antigen pre-S2 (middle envelope protein)
 C:Species: hepatitis B virus, HBV
 A:Variety: subtype ayw, patient A
 C:Date: 06-Jan-1995 #sequence_revision 01-Nov-1996 #text_change 26-Aug-1999
 C:Accession: S41869
 R:Jai, M.E.; Mazzoleni, A.P.; Balestrieri, A.
 Submitted to the EMBL Data Library, January 1994
 A:Description: Sequence analysis of HBV genomes isolated from patients with HBsAg chronic hepatitis B
 A:Reference number: S41869
 A:Accession: S41869
 A:Molecule type: DNA
 A:Residues: 1378 <LAI>
 A:Cross-references: EMBL:X77309; NID:G452611; PIDN:CAA54515.1; PID:G452612
 A:Experimental source: subtype ayw, patient A
 C:Genetics:
 A:Gene: S
 C:Superfamily: hepatitis B virus surface antigen
 C:Keywords: surface antigen
 F:1-378/Product: surface antigen pre-S1 (large envelope protein) #status predicted <PS1>
 F:1-108/Domain: pre-S1 domain #status predicted <PRB1>
 F:109-378/Product: surface antigen pre-S2 (middle envelope protein) #status predicted <PS2>
 F:109-163/Domain: pre-S2 domain #status predicted <PRB2>
 F:164-378/Product: surface antigen S (small envelope protein) #status predicted <PSD>

Query Match 93.3%; Score 1799; DB 2; Length 378;
 Best Local Similarity 97.9%; Pred. No. 5.9e-121;
 Matches 324; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 6 GAGAFGLGFTPPHGGILGWSPOAGILLETLPANPPASTNRSGROPTPLSPPLANTHPQ 65
 DB 48 GAGAFGLGFTPPHGGILGWSPOAGILLETLPANPPASTNRSGROPTPLSPPLANTHPQ 107
 QY 66 AMQNMSTTFHOTLQDPRVRGLYFPAGSSSGTVNPVTVPISISIFSRIGDPALNMENI 125
 DB 108 AMQNMSTTFHOTLQDPRVRGLYFPAGSSSGTVNPVTVPISISIFSRIGDPALNMENI 167
 QY 126 TSGFGLPLVLOAGFFLLTRILITPQSLDSWMTSLNPLGCTTVCLGQNSQSPSNNHSPTS 185
 DB 168 TSGFGLPLVLOAGFFLLTRILITPQSLDSWMTSLNPLGCTTVCLGQNSQSPSNNHSPTS 227
 QY 186 CPPTCGYRMWCLRRFFILFLILLCIFLVLVDYQMLPVCPLIPGSSSTTSGPCRTC 245
 DB 228 CPPTCGYRMWCLRRFFILFLILLCIFLVLVDYQMLPVCPLIPGSSSTTSGPCRTC 287
 QY 246 TTPAGSTMYPSGCCCTKPSDNCCTCIPSSWAFGKFLMWSARFWSLSLVFVQMFV 305
 DB 288 TTPAGSTMYPSGCCCTKPSDNCCTCIPSSWAFGKFLMWSARFWSLSLVFVQMFV 347
 QY 306 GLSPTVWLSVIMMMWYMGPSLYSLISPLPLIPFCCLMVTI 336
 DB 348 GLSPTVWLSVIMMMWYMGPSLYSLISPLPLIPFCCLMVTI 378

RESULT 11
 S20749
 surface antigen - hepatitis B virus (subtype ayw, patient CI)
 N:Alternate names: envelope protein; HBs antigen
 N:Contains: surface antigen pre-S1 (large envelope protein); surface antigen S (small envelope protein)
 C:Species: hepatitis B virus, HBV
 A:Variety: subtype ayw, patient CI
 C:Date: 20-Feb-1995 #sequence_revision 01-Nov-1996 #text_change 26-Aug-1999
 C:Accession: S20749
 R:Jai, M.E.; Mazzoleni, A.P.; Balestrieri, A.; Melis, A.; Porru, A.
 Submitted to the EMBL Data Library, March 1992
 A:Description: Sequence analysis of HBV genomes isolated from patients with HBsAg negative chronic hepatitis B
 A:Reference number: S20745
 A:Accession: S20745
 A:Molecule type: DNA
 A:Residues: 1389 <LAI>
 A:Cross-references: EMBL:X65258; NID:G59434; PIDN:CAA46353.1; PID:G59435
 A:Experimental source: subtype ayw, patient CI
 A:Note: typical methionine for the beginning of pre-S2 domain at position 109 is missing
 C:Genetics:
 A:Gene: S
 C:Superfamily: hepatitis B virus surface antigen
 C:Keywords: surface antigen
 F:1-389/Product: surface antigen pre-S1 (large envelope protein) #status predicted <PS1>
 F:164-389/Product: surface antigen S (small envelope protein) #status predicted <PSD>

Query Match 93.2%; Score 1797; DB 2; Length 389;
 Best Local Similarity 93.0%; Pred. No. 8.4e-121;
 Matches 319; Conservative 11; Mismatches 13; Indels 0; Gaps 0;

QY 5 DGAGAFGLGFTPPHGGILGWSPOAGILLETLPANPPASTNRSGROPTPLSPPLANTHPQ 64
 DB 47 DGAGAFGLGFTPPHGGILGWSPOAGILLETLPANPPASTNRSGROPTPLSPPLANTHPQ 106
 QY 66 AMQNMSTTFHOTLQDPRVRGLYFPAGSSSGTVNPVTVPISISIFSRIGDPALNMENI 124
 DB 107 AMQNMSTTFHOTLQDPRVRGLYFPAGSSSGTVNPVTVPISISIFSRIGDPALNMENI 166
 QY 126 TSGFGLPLVLOAGFFLLTRILITPQSLDSWMTSLNPLGCTTVCLGQNSQSPSNNHSPTS 184
 DB 167 TSGFGLPLVLOAGFFLLTRILITPQSLDSWMTSLNPLGCTTVCLGQNSQSPSNNHSPTS 226
 QY 186 SCPTCGYRMWCLRRFFILFLILLCIFLVLVDYQMLPVCPLIPGSSSTTSGPCRTC 244

Db 227 SCEPTGCGYRMCLRRFTIFLLCLIFLLVLDVQGLHVCPLIPGTTTGTGPKCT 286
 QY 245 CTTPAOCTSMYPSCCCTKPSDGNCTCIPSSNAFGKFLWENASRFSMLSLVPVQMF 304
 Db 287 CTTPAOCTSMYPSCCCTKPSDGNCTCIPSSNAFGKFLWENASRFSMLSLVPVQMF 346
 QY 305 VGLSPYVLSVIMMMYWGSPSLYSILSPFLPLPIFCFLMVIYI 347
 Db 347 VGLSPYVLSVIMMMYWGSPSLYSILSPFLPLPIFCFLMVIYI 389

RESULT 12

SAVLA

large surface antigen - hepatitis B virus (subtype adr)

N:Contains: major surface antigen; middle surface antigen

C:Species: hepatitis B virus, HBV

C:Date: 14-Nov-1983 #sequence revision 14-Nov-1983 #text change 07-May-1999

C:Accession: A03705; S04569; J02107; P00608

R:Ono, Y.; Ono, H.; Saeda, R.; Igatah, K.; Sugino, Y.; Nishio, K.

Nucleic Acids Res. 11, 1747-1757, 1983

A:Title: The complete nucleotide sequence of the cloned hepatitis B virus DNA; subtype

A:Reference number: A93460; MUID:83168919; PMID:6300776

A:Accession: A03705

A:Molecule type: DNA

A:Residues: 1-400 <NO>

A:Cross-references: GB:V00867

R:Rho, H.M.; Kim, K.; Hyun, S.W.; Kim, Y.S.

Nucleic Acids Res. 17, 2124, 1989

A:Title: The nucleotide sequence and reading frames of a mutant hepatitis B virus subtype

A:Reference number: S04568; MUID:89183619; PMID:2928116

A:Accession: S04569

A:Structure: translation not shown

A:Molecule type: DNA

A:Residues: 1-50, 'R', 52-66, 'VP', 69-129, 'V', 131-142, 'P', 144-164, 'S', 166-176, 'S', 178-338, 'A:Cross-references: EMBL:X14193

R:Norde, H.; Hammas, B.; Lee, S.D.; Bille, K.; Courouce, A.M.; Mushahwar, I.K.; Magnus, J. Gen. Virol. 74, 1341-1348, 1993

A:Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin

A:Reference number: J02044; MUID:93329382; PMID:8336122

A:Accession: J02107

A:Molecule type: DNA

A:Residues: 175-400 <NO>

A:Experimental source: Genogroup C, subtype adr, strain pBRHadr4

R:Norde, H.; Courouce, A.M.; Magnus, L.O.

J. Gen. Virol. 73, 3141-3145, 1992

A:Title: Molecular basis of hepatitis B virus serotype variations within the four major

A:Reference number: P00453; MUID:93107848; PMID:1469353

A:Accession: P00608

A:Molecule type: DNA

A:Residues: 275-354 <NO>

A:Experimental source: subtype adr, Bau

C:Genetics:

A:Gene: pre-S1/pre-S2/S

C:Superfamily: hepatitis B virus surface antigen

C:Keywords: glycoprotein; surface antigen

F:110-400/Product: middle surface antigen (gene pre-S2/S) #status predicted <DSA>

F:115-400/Product: major surface antigen (gene S) #status predicted <MSA>

F:115-123/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 91.8%; Score 1771; DB 1; Length 400;
 Best Local Similarity 92.1%; Pred. No. 6,1e-119;
 Matches 315; Conservative 10; Mismatches 17; Indels 0; Gaps 0;

QY 6 GAGAGLGTTPHGLGMSPOAGILTLTPANPPASTNROSGRQPTPLSPPLANTHPQ 65
 DB 59 GAGAGPGGTPPHGGLGMSPOAGILTLTPANPPASTNROSGRQPTPLSPPLANTHPQ 118
 QY 66 ANOMNSTTHQTLQDPRVGLYFPAGSSSGTVNVPVTPVSPISIFSRIGDPAIMENI 125
 DB 119 ANOMNSTTHQTLQDPRVGLYFPAGSSSGTVNVPVTPVSPISIFSRIGDPAIMENI 178
 QY 126 TSGFLPBLVLQAGFLTLRLITPQSLDWSMTSLNPLGCTTVCCGONSQSTSHSPS 185

Db 179 TSGFLPBLVLQAGFLTLRLITPQSLDWSMTSLNPLGAPTCPCGONSQSTSHSPS 238
 QY 186 CPTTCGGRMCLRRFTIFLLCLIFLLVLDVQGLHVCPLIPGTTTGTGPKCTC 245
 Db 229 CPTTCGGRMCLRRFTIFLLCLIFLLVLDVQGLHVCPLIPGTTTGTGPKCTC 298
 QY 246 TTPAOCSTMYPSCCCTKPSDGNCTCIPSSNAFGKFLWENASRFSMLSLVPVQMF 305
 Db 299 TTPAOCSTMYPSCCCTKPSDGNCTCIPSSNAFGKFLWENASRFSMLSLVPVQMF 358
 QY 306 VGLSPYVLSVIMMMYWGSPSLYSILSPFLPLPIFCFLMVIYI 347
 Db 359 VGLSPYVLSVIMMMYWGSPSLYSILSPFLPLPIFCFLMVIYI 400

RESULT 13

SAVLCF

large surface antigen - hepatitis B virus

N:Contains: major surface antigen; middle surface antigen

C:Species: hepatitis B virus, HBV

C:Date: 30-Jun-1989 #sequence revision 30-Jun-1989 #text change 16-Jun-2000

C:Accession: C28885; J02123

R:Vaudin, M.; Wolstenholme, A.J.; Tsiquaye, K.N.; Zuckerman, A.J.; Harrison, T.J.

J. Gen. Virol. 69, 1383-1389, 1988

A:Title: The complete nucleotide sequence of the genome of a hepatitis B virus isolated ;

A:Reference number: A92796; MUID:88258473; PMID:2838576

A:Accession: C28885

A:Molecule type: DNA

A:Residues: 1-389 <NA>

A:Cross-references: GB:D00220; NID:9221505; PIDN:BA00159.1; PID:9221508

A:Experimental source: strain LSH, chimpanzee

R:Norde, H.; Hammas, B.; Lee, S.D.; Bille, K.; Courouce, A.M.; Mushahwar, I.K.; Magnus, J. Gen. Virol. 74, 1341-1348, 1993

A:Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin

A:Reference number: J02044; MUID:93329382; PMID:8336122

A:Accession: J02123

A:Molecule type: DNA

A:Residues: 164-389 <NO>

A:Experimental source: subtype adr2, strain adrLSH

C:Genetics:

A:Gene: pre-S1/pre-S2/S; S

C:Superfamily: hepatitis B virus surface antigen

C:Keywords: glycoprotein; surface antigen

F:103-388/Product: middle surface antigen (gene pre-S2/S) #status predicted <DSA>

F:164-389/Product: major surface antigen (gene S) #status predicted <MSA>

F:112-166/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 91.0%; Score 1755; DB 1; Length 389;
 Best Local Similarity 91.5%; Pred. No. 8,2e-118;
 Matches 313; Conservative 9; Mismatches 20; Indels 0; Gaps 0;

QY 6 GAGAGLGTTPHGLGMSPOAGILTLTPANPPASTNROSGRQPTPLSPPLANTHPQ 65
 Db 48 GAGAGPGGTPPHGGLGMSPOAGILTLTPANPPASTNROSGRQPTPLSPPLANTHPQ 107
 QY 66 ANOMNSTTHQTLQDPRVGLYFPAGSSSGTVNVPVTPVSPISIFSRIGDPAIMENI 125
 Db 108 ANOMNSTTHQTLQDPRVGLYFPAGSSSGTVNVPVTPVSPISIFSRIGDPAIMENI 167
 QY 126 TSGFLPBLVLQAGFLTLRLITPQSLDWSMTSLNPLGCTTVCCGONSQSTSHSPS 185
 Db 168 TSGFLPBLVLQAGFLTLRLITPQSLDWSMTSLNPLGAPTCPCGONSQSTSHSPS 227
 QY 186 CPTTCGGRMCLRRFTIFLLCLIFLLVLDVQGLHVCPLIPGTTTGTGPKCTC 245
 Db 228 CPTTCGGRMCLRRFTIFLLCLIFLLVLDVQGLHVCPLIPGTTTGTGPKCTC 287
 QY 246 TTPAOCSTMYPSCCCTKPSDGNCTCIPSSNAFGKFLWENASRFSMLSLVPVQMF 305
 Db 288 TTPAOCSTMYPSCCCTKPSDGNCTCIPSSNAFGKFLWENASRFSMLSLVPVQMF 347
 QY 306 VGLSPYVLSVIMMMYWGSPSLYSILSPFLPLPIFCFLMVIYI 347

Db 348 GLSPTVWLSAIIWMWYMGPSLYISILSPFLPLIFPCIMVYI 389

RESULT 14

S35528

surface antigen - hepatitis B virus (subtype adr)

N/Alternate names: envelope protein; HBs antigen

C/Species: hepatitis B virus, HBV

A/Variety: subtype adr

C/Date: 09-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 20-Jun-2000

C/Accession: S35528

R/Mukaido, M.; Kumazawa, T.; Hoshi, A.; Kawaguchi, R.; Hiki, K.

Nucleic Acids Res. 20, 6105, 1992

A/Title: The complete nucleotide sequence of hepatitis B virus, subtype adr (SRADR) and

A/Reference number: S35527; MUID:9309607; PMID:1461746

A/Accession: S35528

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-400 <NR>

A/Cross-references: EMBL:DJ2980; NID:9221500; PDB:BA02359.1; PID:g221502

A/Experimental source: subtype adr

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1992

C/Genetics:

A:Gene: S

A:Introns: 165/2

C/Superfamily: hepatitis B virus surface antigen

C/Keywords: surface antigen

F:1-400/Product: surface antigen pre-S1 (large envelope protein) #status predicted <PS1>

F:1-119/Domain: pre-S1 domain #status predicted <PRE1>

F:120-400/Product: surface antigen pre-S2 (middle envelope protein) #status predicted <PS2>

F:175-400/Product: surface antigen S (small envelope protein) #status predicted <PSD>

Query Match

Best Local Similarity 90.5%; Score 1745; DB 2; Length 400;

Matches 311; Conservative 11; Mismatches 20; Indels 0; Gaps 0;

6 GAGAFGLGFTPHGGLGMSPOAGILETLPANPPASTNRSGROPTPLSPRLANTHPQ 65
 59 GAGDGFPGFTPHGGLGMSPOAGILETLTPAPAPASTNRSGROPTPLSPRLANTHPQ 118

66 AMONNSTTFHQTLODPRVRGLYFPAGSSSGTVNPVTTSPISSISFRIGDPAIAMENTI 125
 119 AMONNSTTFHQTLODPRVRGLYFPAGSSSGTVNPVTTSPISSISFRIGDPAIAMENTI 178

126 TSGFGLPLVLOAGFFLLTRILITPQSLDSWMTSLNPLGAGTCGQNSQSPHSPTS 185
 179 TSGFGLPLVLOAGFFLLTRILITPQSLDSWMTSLNPLGAGTCGQNSQSPHSPTS 238

186 CPPTCGYRMWCLRRFIIIFLLICLIFLLVLDYQMLPVCPLIPGSSSTTGPCTC 245
 239 CPPTCGYRMWCLRRFIIIFLLICLIFLLVLDYQMLPVCPLIPGSSSTTGPCTC 298

246 TTPAGTSMYPSCCCTKPSDNCCTCIPSSWAFKFLMWSARFSLSLVPEVQMFV 305
 299 TTPAGTSMYPSCCCTKPSDNCCTCIPSSWAFKFLMWSARFSLSLVPEVQMFV 358

306 GLSPTVWLSAIIWMWYMGPSLYISILSPFLPLIFPCIMVYI 347
 359 GLSPTVWLSAIIWMWYMGPSLYISILSPFLPLIFPCIMVYI 400

RESULT 15

S41870

surface antigen - hepatitis B virus (subtype ayw, patient M) (fragment)

N/Alternate names: envelope protein; HBs antigen

C/Species: hepatitis B virus, HBV

A/Variety: subtype ayw, patient M

C/Date: 06-Jan-1995 #sequence_revision 01-Nov-1996 #text_change 26-Aug-1999

C/Accession: S41870

R/Lai, M.E.; Mazzoleni, A.P.; Balestrieri, A.

submitted to the EMBL Data Library, January 1994

A/Description: Sequence analysis of HBV genomes isolated from patients with HBsAg chronic

A/Reference number: S41869

A/Accession: S41870

A/Molecule type: DNA

A/Residues: 1-378 <LAI>

A/Cross-references: EMBL:X77308; NID:9452613; PDB:CMAS4514.1; PID:9452614

A/Experimental source: subtype ayw, patient M

C/Genetics:

A:Gene: S

C/Superfamily: hepatitis B virus surface antigen

C/Keywords: surface antigen

F:1-378/Product: surface antigen pre-S1 (large envelope protein) #status predicted <PS1>

F:1-108/Domain: pre-S1 domain #status predicted <PRE1>

F:109-378/Product: surface antigen pre-S2 (middle envelope protein) #status predicted <PS2>

F:164-378/Product: surface antigen S (small envelope protein) #status predicted <PSD>

Query Match

Best Local Similarity 89.9%; Score 1734; DB 2; Length 378;

Matches 311; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

6 GAGAFGLGFTPHGGLGMSPOAGILETLPANPPASTNRSGROPTPLSPRLANTHPQ 65
 48 GAGAFGLGFTPHGGLGMSPOAGILETLTPANPPASTNRSGROPTPLSPRLANTHPQ 107

66 AMONNSTTFHQTLODPRVRGLYFPAGSSSGTVNPVTTSPISSISFRIGDPAIAMENTI 125
 108 AMONNSTTFHQTLODPRVRGLYFPAGSSSGTVNPVTTSPISSISFRIGDPAIAMENTI 167

126 TSGFGLPLVLOAGFFLLTRILITPQSLDSWMTSLNPLGAGTCGQNSQSPHSPTS 185
 168 TSGFGLPLVLOAGFFLLTRILITPQSLDSWMTSLNPLGAGTCGQNSQSPHSPTS 227

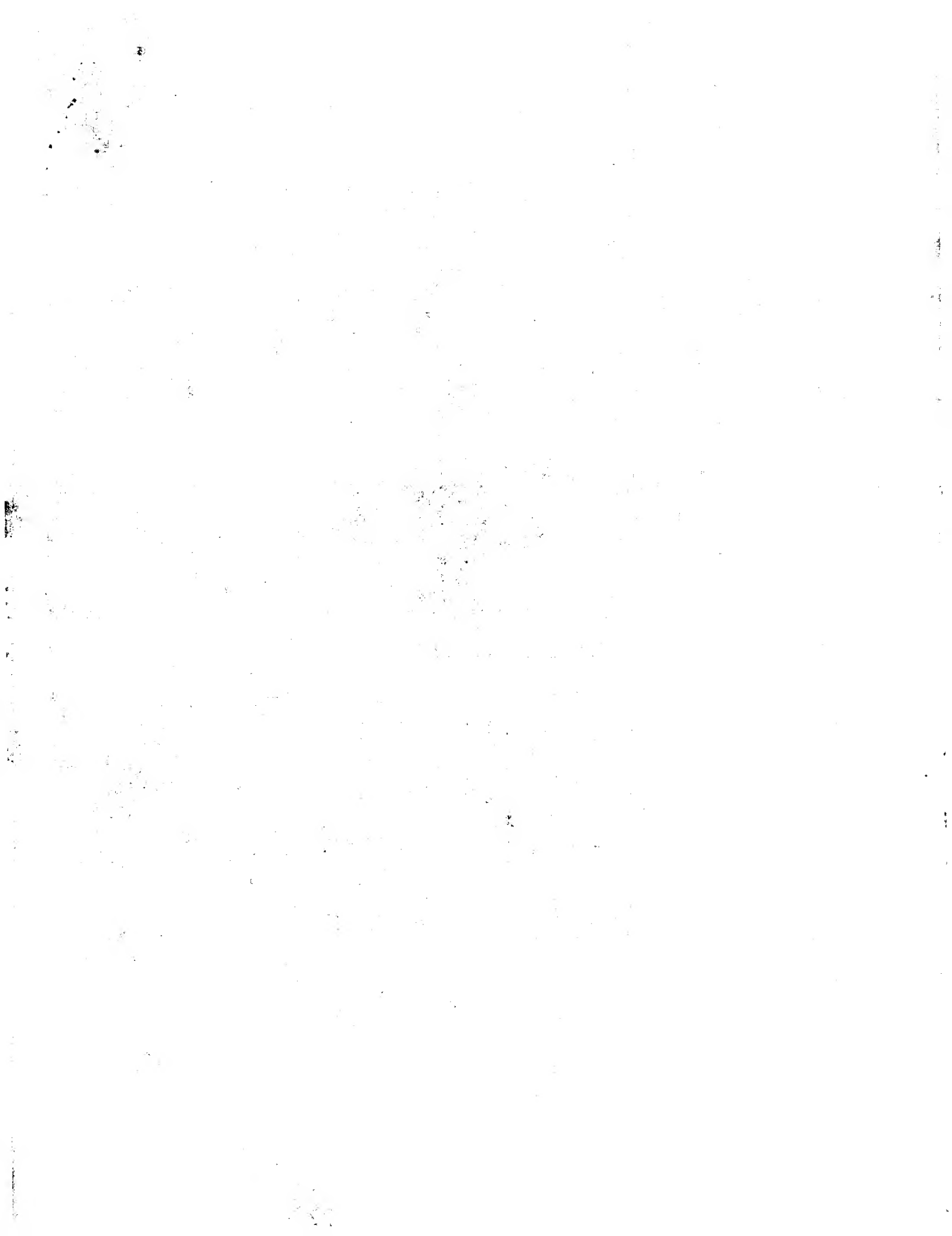
186 CPPTCGYRMWCLRRFIIIFLLICLIFLLVLDYQMLPVCPLIPGSSSTTGPCTC 245
 228 CPPTCGYRMWCLRRFIIIFLLICLIFLLVLDYQMLPVCPLIPGSSSTTGPCTC 287

246 TTPAGTSMYPSCCCTKPSDNCCTCIPSSWAFKFLMWSARFSLSLVPEVQMFV 305
 288 TTPAGTSMYPSCCCTKPSDNCCTCIPSSWAFKFLMWSARFSLSLVPEVQMFV 347

306 GLSPTVWLSAIIWMWYMGPSLYISILSPFLPLIFPCIMVYI 336
 348 GLSPTVWLSAIIWMWYMGPSLYISILSPFLPLIFPCIMVYI 378

Search completed: February 3, 2003, 09:26:27

Job time: 23.9721 sec



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 09:24:27 ; Search time 28.4129 Seconds

(Without alignments)
988,252 Million cell updates/sec

Title: US-09-890-752a-1

Perfect score: 1929
Sequence: 1 MGGDAGAFGLGFTPHGSG.....SLSPFLPLPLTFTCLWYI 347

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 404799 seqs, 80919614 residues

Total number of hits satisfying chosen parameters: 404799

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Pending Patents AA New:*
1: /cgn2_6/pdata/1/paa/PCT_NEM_COMB.pep:*
2: /cgn2_6/pdata/1/paa/US06_NEM_COMB.pep:*
3: /cgn2_6/pdata/1/paa/US07_NEM_COMB.pep:*
4: /cgn2_6/pdata/1/paa/US08_NEM_COMB.pep:*
5: /cgn2_6/pdata/1/paa/US09_NEM_COMB.pep:*
6: /cgn2_6/pdata/1/paa/US10_NEM_COMB.pep:*
7: /cgn2_6/pdata/1/paa/US60_NEM_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1756	91.0	400	1	PCT-US02-32263-46
2	1756	91.0	400	6	US-10-287-994-46
3	300	15.6	67	6	US-10-326-908-21
4	298	15.4	64	6	US-10-326-908-19
5	131	6.8	63	6	US-10-326-908-18
6	124	6.4	68	6	US-10-326-908-20
7	107	5.5	20	5	US-09-863-054-22
8	105	5.4	276	5	US-09-724-676-50955
9	105	5.4	276	5	US-09-724-676-50956
10	105	5.4	276	5	US-09-724-676-50957
11	105	5.4	276	5	US-09-724-676-50958
12	105	5.4	276	5	US-09-724-676-50959
13	105	5.4	276	5	US-09-724-676-50960
14	105	5.4	328	5	US-09-724-676-50953
15	105	5.4	328	5	US-09-724-676-50954
16	105	5.4	328	5	US-09-724-676-50955
17	105	5.4	328	5	US-09-724-676-50956
18	103	5.3	19	5	US-09-863-054-28
19	102	5.3	682	6	US-10-144-779-327
20	101.5	5.3	2509	6	US-10-237-271-1
21	100.5	5.2	641	6	US-10-311-034-15
22	99.5	5.2	845	6	US-09-724-676-94342
23	99.5	5.2	845	6	US-09-724-676-94342
24	99.5	5.2	1078	6	US-10-170-385-295
25	99.5	5.2	1088	6	US-10-284-130-13
26	99	5.1	18	5	US-09-863-054-29

27	97	5.0	1251	1	PCT-US02-10780-80	Sequence 80, Appl
28	97	5.0	1251	6	US-10-114-270-80	Sequence 80, Appl
29	96	5.0	611	6	US-10-281-346-4	Sequence 4, Appl
30	95.5	5.0	1367	1	PCT-US02-40225-3355	Sequence 3355, Ap
31	95.5	5.0	1367	6	US-10-320-797-3355	Sequence 3355, Ap
32	95.5	5.0	1441	6	US-10-170-682-3	Sequence 30, Appl
33	95	4.9	17	5	US-09-863-054-20	Sequence 86312, A
34	95	4.9	606	5	US-09-724-676-86512	Sequence 86512, A
35	95	4.9	606	5	US-09-724-676-86513	Sequence 86513, A
36	95	4.9	606	5	US-09-724-676-86514	Sequence 86514, A
37	95	4.9	606	5	US-09-724-676-86515	Sequence 86515, A
38	95	4.9	606	5	US-09-724-676-86516	Sequence 86516, A
39	95	4.9	606	5	US-09-724-676-86517	Sequence 86517, A
40	95	4.9	606	5	US-09-724-676-86518	Sequence 86518, A
41	95	4.9	606	5	US-09-724-676-86519	Sequence 86519, A
42	95	4.9	798	5	US-09-724-676-86525	Sequence 86525, A
43	95	4.9	798	5	US-09-724-676-86526	Sequence 86526, A
44	95	4.9	798	5	US-09-724-676-86527	Sequence 86527, A
45	95	4.9	798	5	US-09-724-676-86528	Sequence 86528, A

ALIGNMENTS

```
RESULT 1
PCT-US02-32263-46
; Sequence 46, Application PC/TUS0232263
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: Defrees, Shawn
; APPLICANT: Zopf, David
; APPLICANT: Bayfer, Robert
; APPLICANT: Bove, Caryn
; APPLICANT: Hakes, David
; APPLICANT: Chen, Xi
; TITLE OF INVENTION: REMODELING AND GLYCOCONJUGATION OF PEPTIDES
; FILE REFERENCE: 040853-01-5050MO
; CURRENT APPLICATION NUMBER: PCT/US02/32263
; PRIOR FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US 60/328,523
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/344,692
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/334,233
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/334,301
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/387,292
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 60/391,777
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46
; LENGTH: 400
; TYPE: PRT
; ORGANISM: Hepatitis B virus
PCT-US02-32263-46
```

Query Match: 91.0%; Score 1756; DB 1; Length 400;
Best Local Similarity: 91.2%; Pred. No. 4.5e-106;
Matches 312; Conservative 11; Mismatches 19; Indels 0; Gaps 0;

QY 6 GAGAFGLGFTPHGGLGMSPOAGSLFTLPANPPASTNRSGROPTPLSPPLANTHPQ 65
DB 59 GAGDFGFGFTPHGGLGMSPOAGSLFTLPANPPASTNRSGROPTPLSPPLANDSHPO 118

```
Qy 66 AMONSTTFHOTLDDPRVRLGYPAGSSSGTVNVPPTVSPISISFRIGDPALNMENI 125
Db 119 AMONSTTFHOTLDDPRVRLGYPAGSSSGTVNVPPTVSPISISFRIGDPALNMENI 178
Qy 126 TSGFLGPLLVLAQAGFLLRLITLTPQSLDSWMTSLNPLAGTTVCLQONSQSPSTNSHPTS 185
Db 179 TSGFLGPLLVLAQAGFLLRLITLTPQSLDSWMTSLNPLAGTTVCLQONSQSPSTNSHPTS 238
Qy 186 CEPICGVRMCLRRPFIIFLLCLIFLLVLDYQGLPVCPILPGSSTTSGPCRTC 245
Db 239 CEPICGVRMCLRRPFIIFLLCLIFLLVLDYQGLPVCPILPGSSTTSGPCRTC 298
Qy 246 TTPAGTSMVPSCCCTKPSDGNCTCIPSSNAFGKFLWEMASARFWSLILVFPVQWFA 305
Db 299 TTPAGTSMVPSCCCTKPSDGNCTCIPSSNAFGKFLWEMASARFWSLILVFPVQWFA 358
Qy 306 GLSPTVWLSVIMMMWYMGPSLYSILSPFLPLPIFFCLMWYI 347
Db 359 GLSPTVWLSVIMMMWYMGPSLYSILSPFLPLPIFFCLMWYI 400

RESULT 2
US-10-287-994-46
; Sequence 46, Application US/10287994
; GENERAL INFORMATION:
; APPLICANT: Nectose Technologies, Inc.
; APPLICANT: Defreese, Shawn
; APPLICANT: Zopf, David
; APPLICANT: Bayer, Robert
; APPLICANT: Bower, Caryn
; APPLICANT: Hakes, David
; APPLICANT: Chen, Xi
; TITLE OF INVENTION: REMODELING AND GLYCOCONJUGATION OF PEPTIDES
; FILE REFERENCE: 040853-01-5052-00
; CURRENT FILING DATE: US/10/287,994
; PRIOR APPLICATION NUMBER: US 60/328,523
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/344,692
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/387,292
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 60/391,777
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46
; LENGTH: 400
; TYPE: PRT
; ORGANISM: Hepatitis B virus
US-10-287-994-46

Query Match 91.0%; Score 1756; DB 6; Length 400;
Best Local Similarity 91.2%; Pred. No. 4,5e-106;
Matches 312; Conservative 11; Mismatches 19; Indels 0; Gaps 0;
```

```
Qy 186 CEPICGVRMCLRRPFIIFLLCLIFLLVLDYQGLPVCPILPGSSTTSGPCRTC 245
Db 239 CEPICGVRMCLRRPFIIFLLCLIFLLVLDYQGLPVCPILPGSSTTSGPCRTC 298
Qy 246 TTPAGTSMVPSCCCTKPSDGNCTCIPSSNAFGKFLWEMASARFWSLILVFPVQWFA 305
Db 299 TTPAGTSMVPSCCCTKPSDGNCTCIPSSNAFGKFLWEMASARFWSLILVFPVQWFA 358
Qy 306 GLSPTVWLSVIMMMWYMGPSLYSILSPFLPLPIFFCLMWYI 347
Db 359 GLSPTVWLSVIMMMWYMGPSLYSILSPFLPLPIFFCLMWYI 400

RESULT 3
US-10-326-908-21
; Sequence 21, Application US/10326908
; GENERAL INFORMATION:
; APPLICANT: Christopher JONES
; APPLICANT: Andrew BACON
; APPLICANT: Gill DOUCE
; APPLICANT: Mark PAGE
; TITLE OF INVENTION: Designing Immunogens
; FILE REFERENCE: 117-433 / N78027D
; CURRENT FILING DATE: US/10/326,908
; PRIOR APPLICATION NUMBER: US 10/110,036
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: PCT/GB00/03857
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: GB 0007789.1
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: GB 9923902.2
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: MS Word
; SEQ ID NO 21
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: synthetic peptide immunogen
US-10-326-908-21

Query Match 15.6%; Score 300; DB 6; Length 67;
Best Local Similarity 100.0%; Pred. No. 8,3e-13;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 65 QAMONSTTFHOTLDDPRVRLGYPAGSSSGTVNVPPTVSPISISFRIGDPALN 121
Db 6 QAMONSTTFHOTLDDPRVRLGYPAGSSSGTVNVPPTVSPISISFRIGDPALN 62

RESULT 4
US-10-326-908-19
; Sequence 19, Application US/10326908
; GENERAL INFORMATION:
; APPLICANT: Christopher JONES
; APPLICANT: Andrew BACON
; APPLICANT: Gill DOUCE
; APPLICANT: Mark PAGE
; TITLE OF INVENTION: Designing Immunogens
; FILE REFERENCE: 117-433 / N78027D
; CURRENT FILING DATE: US/10/326,908
; PRIOR APPLICATION NUMBER: US 10/110,036
; PRIOR FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: US 10/110,036
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: PCT/GB00/03857
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: GB 0007789.1
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: GB 9923902.2
; PRIOR FILING DATE: 1999-10-08
```

NUMBER OF SEQ ID NOS: 30
 SOFTWARE: MS Word
 SEQ ID NO 15
 LENGTH: 64
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: synthetic peptide immunogen
 US-10-326-908-19

Query Match 15.4%; Score 298; DB 6; Length 64;
 Best Local Similarity 96.6%; Pred. No. 1.1e-12;
 Matches 57; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 65 GAGMGNSTTHQTIQDRVRGLYPACGSSSGTVNPPVTVSPISSTFSRIGDPALNME 123
 Db 6 GAGMGNSTTHQTIQDRVRGLYPACGSSSGTVNPPVTVSPISSTFSRIGDPALNME 64

RESULT 5
 US-10-326-908-18
 Sequence 19 Application US/10326908
 GENERAL INFORMATION:
 APPLICANT: Christopher JONES
 APPLICANT: Andrew BACON
 APPLICANT: GILL DOUCE
 APPLICANT: Mark PAGE
 TITLE OF INVENTION: Designing Immunogens
 FILE REFERENCE: 117-433 / N78027D
 CURRENT APPLICATION NUMBER: US/10/326,908
 PRIOR FILING DATE: 2002-12-23
 PRIOR APPLICATION NUMBER: US 10/110,036
 PRIOR FILING DATE: 2002-04-08
 PRIOR APPLICATION NUMBER: PCT/GB00/03857
 PRIOR FILING DATE: 2000-10-06
 PRIOR APPLICATION NUMBER: GB 0007789.1
 PRIOR FILING DATE: 2000-03-30
 PRIOR APPLICATION NUMBER: GB 9923902.2
 NUMBER OF SEQ ID NOS: 30
 SOFTWARE: MS Word
 SEQ ID NO 18
 LENGTH: 63
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: synthetic peptide immunogen
 US-10-326-908-18

Query Match 6.8%; Score 131; DB 6; Length 63;
 Best Local Similarity 91.7%; Pred. No. 0.064;
 Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GAGFGLGFTPHGGLGMSPOAG 29
 Db 40 GAGFGLGFTPHGGLGMSPOSE 63

RESULT 6
 US-10-326-908-20
 Sequence 20 Application US/10326908
 GENERAL INFORMATION:
 APPLICANT: Christopher JONES
 APPLICANT: Andrew BACON
 APPLICANT: GILL DOUCE
 APPLICANT: Mark PAGE
 TITLE OF INVENTION: Designing Immunogens
 FILE REFERENCE: 117-433 / N78027D
 CURRENT APPLICATION NUMBER: US/10/326,908
 PRIOR FILING DATE: 2002-12-23
 PRIOR APPLICATION NUMBER: US 10/110,036
 PRIOR FILING DATE: 2002-04-08
 PRIOR APPLICATION NUMBER: PCT/GB00/03857

PRIOR FILING DATE: 2000-10-06
 PRIOR APPLICATION NUMBER: GB 0007789.1
 PRIOR FILING DATE: 2000-03-30
 PRIOR APPLICATION NUMBER: GB 9923902.2
 PRIOR FILING DATE: 1999-10-08
 NUMBER OF SEQ ID NOS: 30
 SOFTWARE: MS Word
 SEQ ID NO 20
 LENGTH: 68
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: synthetic peptide immunogen
 US-10-326-908-20

Query Match 6.4%; Score 124; DB 6; Length 68;
 Best Local Similarity 88.0%; Pred. No. 0.19;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 6 GAGFGLGFTPHGGLGMSPOAG 30
 Db 40 GAGFGLGFTPHGGLGMSPHDIG 64

RESULT 7
 US-09-863-054-22
 Sequence 22 Application US/09863054
 GENERAL INFORMATION:
 APPLICANT: Chisari, Francis V.
 TITLE OF INVENTION: Peptides for Inducing Cytotoxic T Lymphocyte Responses to Hepatitis B Virus
 NUMBER OF SEQUENCES: 57
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Townsend and Townsend and Crew LLP
 STREET: Two Bimbarcadero Center, Eighth Floor
 CITY: San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94111-3834
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/863,054
 FILING DATE: 21-May-2001
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/935,898
 FILING DATE: 26-AUG-1992
 APPLICATION NUMBER: US 08/024,120
 FILING DATE: 26-FEB-1993
 APPLICATION NUMBER: US 08/396,283
 FILING DATE: 27-FEB-1995
 APPLICATION NUMBER: US 08/463,486
 FILING DATE: 05-JUN-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Lockyer, Jean M.
 REGISTRATION NUMBER: 44,879
 REFERENCE/DOCKET NUMBER: 014740-000421US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 576-0200
 TELEFAX: (415) 576-0300
 INFORMATION FOR SEQ ID NO: 22:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 20 amino acids
 TYPE: amino acid
 STRANDEDNESS: <Unknown>
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 SEQUENCE DESCRIPTION: SEQ ID NO: 22:
 US-09-863-054-22

Query Match 5.5%; Score 107; DB 5; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0.8; Indels 0; Gaps 0;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 287 ASARPSMLLVPEVQWVG 306
 |||||
 DB 1 ASARPSMLLVPEVQWVG 20

RESULT 8
 US-09-724-676-50955
 ; Sequence 50955, Application US/09724676
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724,676
 ; CURRENT FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 50955
 ; LENGTH: 276
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-724-676-50955

Query Match 5.4%; Score 105; DB 5; Length 276;
 Best Local Similarity 26.0%; Pred. No. 12;
 Matches 57; Conservative 11; Mismatches 73; Indels 78; Gaps 10;

QY 2 GRGDGAGAFGLGTPPHGGL-----GMSPOAGILETLPANPP--- 40
 DB 32 GGGGGGSGGGGPPGLGGLFOAGMPKLRSTANRDNDSGSRP-----FLPPGGR 82
 QY 41 -----PASTRQSGROPTPLSPPLRNTHPQAMQNSTTHQTLQDPRVGLYFPAGSSS 95
 DB 83 STSAKPPSPSGRPFVP--SPGHRSGPPEPQ-----NRMPPR-----PDVGSKP 128
 QY 96 GTV-NPVTTPVSPISIFSRIGDPALMENTSGFLGPLVLVQAGFLLRLITIPQSLD 154
 DB 129 DSIPPPVSTPRPIQSSLNRGSP-----VPGG-----PROPS 162

QY 155 SMTSLNPLGTTVCLGONS--QSPTSNSHPTSCPTCP 191
 DB 163 PGTPPPPPGNRGTALGGGSIROSLSSSPFSNRPLP 201

RESULT 9
 US-09-724-676-50956
 ; Sequence 50956, Application US/09724676
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724,676
 ; CURRENT FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 50956
 ; LENGTH: 276
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-724-676-50956

Query Match 5.4%; Score 105; DB 5; Length 276;
 Best Local Similarity 26.0%; Pred. No. 12;
 Matches 57; Conservative 11; Mismatches 73; Indels 78; Gaps 10;

QY 2 GRGDGAGAFGLGTPPHGGL-----GMSPOAGILETLPANPP--- 40
 DB 32 GGGGGGSGGGGPPGLGGLFOAGMPKLRSTANRDNDSGSRP-----FLPPGGR 82

QY 41 -----PASTRQSGROPTPLSPPLRNTHPQAMQNSTTHQTLQDPRVGLYFPAGSSS 95
 DB 83 STSAKPPSPSGRPFVP--SPGHRSGPPEPQ-----NRMPPR-----PDVGSKP 128
 QY 96 GTV-NPVTTPVSPISIFSRIGDPALMENTSGFLGPLVLVQAGFLLRLITIPQSLD 154
 DB 129 DSIPPPVSTPRPIQSSLNRGSP-----VPGG-----PROPS 162

QY 155 SMTSLNPLGTTVCLGONS--QSPTSNSHPTSCPTCP 191
 DB 163 PGTPPPPPGNRGTALGGGSIROSLSSSPFSNRPLP 201

RESULT 10
 US-09-724-676-50957
 ; Sequence 50957, Application US/09724676
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724,676
 ; CURRENT FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 50957
 ; LENGTH: 276
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-724-676-50957

Query Match 5.4%; Score 105; DB 5; Length 276;
 Best Local Similarity 26.0%; Pred. No. 12;
 Matches 57; Conservative 11; Mismatches 73; Indels 78; Gaps 10;

QY 2 GRGDGAGAFGLGTPPHGGL-----GMSPOAGILETLPANPP--- 40
 DB 32 GGGGGGSGGGGPPGLGGLFOAGMPKLRSTANRDNDSGSRP-----FLPPGGR 82
 QY 41 -----PASTRQSGROPTPLSPPLRNTHPQAMQNSTTHQTLQDPRVGLYFPAGSSS 95
 DB 83 STSAKPPSPSGRPFVP--SPGHRSGPPEPQ-----NRMPPR-----PDVGSKP 128
 QY 96 GTV-NPVTTPVSPISIFSRIGDPALMENTSGFLGPLVLVQAGFLLRLITIPQSLD 154
 DB 129 DSIPPPVSTPRPIQSSLNRGSP-----VPGG-----PROPS 162

QY 155 SMTSLNPLGTTVCLGONS--QSPTSNSHPTSCPTCP 191
 DB 163 PGTPPPPPGNRGTALGGGSIROSLSSSPFSNRPLP 201

RESULT 11
 US-09-724-676A-50955
 ; Sequence 50955, Application US/09724676A
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724,676A
 ; CURRENT FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 50955
 ; LENGTH: 276
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-724-676A-50955

Query Match 5.4%; Score 105; DB 5; Length 276;
 Best Local Similarity 26.0%; Pred. No. 12;
 Matches 57; Conservative 11; Mismatches 73; Indels 78; Gaps 10;

QY 2 GRGDGAGAFGLGTPPHGGL-----GMSPOAGILETLPANPP--- 40

```

Db      32 GGGGGGGGFCGGGPPGGLGGLFQAGMPKLRSTANRDNDSGSRP-----PLLPGGR 82
QY      41 -----PASTNROSGROPTPLSPPLRNTHPQAMQMNSTTFHQTLODPRVRLGYPAGSSS 95
Db      83 STSAKFPSPSGRPFVP-SPGHRSGPPEPQ-----NRMPPR-----PDVGSKP 128
QY      96 GTV-NPVPTTSPISISIFSRIGDPALNMENITSGFLPLVLQAGFFLLTRILITPOSID 154
Db      129 DSIPPEVSTPRPIQSSLNHRGSP-----VPGG-----PRQPS 162
QY      155 SWTSLNPLGCTTVCIGONS--QSPTSNSHSPSCPTCP 191
Db      163 PGTPPPFPNGRGTALGGGSIROSPLSSSPFSNRPPLP 201

```

```

RESULT 12
US-09-724-676A-50956
; Sequence 50956, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; NUMBER OF SEQ ID NOS: 2000-11-28
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 50956
; LENGTH: 276
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-50956

```

```

Query Match
Best Local Similarity 5.4%; Score 105; DB 5; Length 276;
Matches 57; Conservative 11; Mismatches 73; Indels 78; Gaps 10;

```

```

QY      2 GRGDGAGAFGLGFTPPHGGIL-----GMSPOAGILETLTPANPP--- 40
Db      32 GGGGGGGGFCGGGPPGGLGGLFQAGMPKLRSTANRDNDSGSRP-----PLLPGGR 82
QY      41 -----PASTNROSGROPTPLSPPLRNTHPQAMQMNSTTFHQTLODPRVRLGYPAGSSS 95
Db      83 STSAKFPSPSGRPFVP-SPGHRSGPPEPQ-----NRMPPR-----PDVGSKP 128
QY      96 GTV-NPVPTTSPISISIFSRIGDPALNMENITSGFLPLVLQAGFFLLTRILITPOSID 154
Db      129 DSIPPEVSTPRPIQSSLNHRGSP-----VPGG-----PRQPS 162
QY      155 SWTSLNPLGCTTVCIGONS--QSPTSNSHSPSCPTCP 191
Db      163 PGTPPPFPNGRGTALGGGSIROSPLSSSPFSNRPPLP 201

```

```

RESULT 13
US-09-724-676A-50957
; Sequence 50957, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; NUMBER OF SEQ ID NOS: 2000-11-28
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 50957
; LENGTH: 276
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-50957

```

```

Query Match
Best Local Similarity 5.4%; Score 105; DB 5; Length 276;
Matches 57; Conservative 11; Mismatches 73; Indels 78; Gaps 10;

```

```

Matches 57; Conservative 11; Mismatches 73; Indels 78; Gaps 10;
QY      2 GRGDGAGAFGLGFTPPHGGIL-----GMSPOAGILETLTPANPP--- 40
Db      32 GGGGGGGGFCGGGPPGGLGGLFQAGMPKLRSTANRDNDSGSRP-----PLLPGGR 82
QY      41 -----PASTNROSGROPTPLSPPLRNTHPQAMQMNSTTFHQTLODPRVRLGYPAGSSS 95
Db      83 STSAKFPSPSGRPFVP-SPGHRSGPPEPQ-----NRMPPR-----PDVGSKP 128
QY      96 GTV-NPVPTTSPISISIFSRIGDPALNMENITSGFLPLVLQAGFFLLTRILITPOSID 154
Db      129 DSIPPEVSTPRPIQSSLNHRGSP-----VPGG-----PRQPS 162
QY      155 SWTSLNPLGCTTVCIGONS--QSPTSNSHSPSCPTCP 191
Db      163 PGTPPPFPNGRGTALGGGSIROSPLSSSPFSNRPPLP 201

```

```

RESULT 14
US-09-724-676-50953
; Sequence 50953, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; NUMBER OF SEQ ID NOS: 2000-11-28
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 50953
; LENGTH: 328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-50953

```

```

Query Match
Best Local Similarity 5.4%; Score 105; DB 5; Length 328;
Matches 57; Conservative 11; Mismatches 73; Indels 78; Gaps 10;

```

```

QY      2 GRGDGAGAFGLGFTPPHGGIL-----GMSPOAGILETLTPANPP--- 40
Db      84 GGGGGGGGFCGGGPPGGLGGLFQAGMPKLRSTANRDNDSGSRP-----PLLPGGR 134
QY      41 -----PASTNROSGROPTPLSPPLRNTHPQAMQMNSTTFHQTLODPRVRLGYPAGSSS 95
Db      135 STSAKFPSPSGRPFVP-SPGHRSGPPEPQ-----NRMPPR-----PDVGSKP 180
QY      96 GTV-NPVPTTSPISISIFSRIGDPALNMENITSGFLPLVLQAGFFLLTRILITPOSID 154
Db      181 DSIPPEVSTPRPIQSSLNHRGSP-----VPGG-----PRQPS 214
QY      155 SWTSLNPLGCTTVCIGONS--QSPTSNSHSPSCPTCP 191
Db      215 PGTPPPFPNGRGTALGGGSIROSPLSSSPFSNRPPLP 253

```

```

RESULT 15
US-09-724-676-50954
; Sequence 50954, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; NUMBER OF SEQ ID NOS: 2000-11-28
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 50954
; LENGTH: 328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-50954

```

Query Match 5.44; Score 105; DB 5; Length 328;
 Best Local Similarity 26.08; Pred. No. 14;
 Matches 57; Conservative 11; Mismatches 73; Indels 78; Gaps 10;

```

QY 2 GRGAGAGFGLGFTPRHGLL-----GWSQAQGIETLPANPP--- 40
DB 84 GGGGGGSGFGGGGPGGLGLFOAGMKLRSTANRDNDSGSRP-----PILPPQGR 134
QY 41 -----PASTNRQSGRQPTPLSEPLRNTHPQAMQWNSSTFHOTLQDPVYRGUYFPAGGSS 95
DB 135 STSAKPFSPSPGPRFVP-SRGRSGPPEPQR-----NRMPEPR-----PDVGSKP 180
QY 96 GTV-NPVPTTSPISISIRIDDPALMENTNSGFLPLVLQAGFLTLRLITIPQSLD 154
DB 181 DSIPPPVPSTPRIOSLHNRGSP-----VPGG-----FRQPS 214
QY 155 SWTSLNFIQTTVCIGONS--QSPTSNHSPSCPTCP 191
DB 215 PGTPPPPPGNRGTALGGSIROSPLSSSPFSNRPPLP 253
  
```

Search completed: February 3, 2003, 09:35:00
 Job time : 30.4129 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 09:23:57 ; Search time 247.857 Seconds

(Without alignments)
902.626 Million cell updates/sec

Title: US-09-890-752A-1

Sequence: 1 MGRGDGAGRGIGFTPHCG.....SLSPFLPLPIFCLMYVI 347

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA Main:*

1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*

2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*

3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*

4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*

5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*

6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*

7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*

8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*

9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*

10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*

11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*

12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*

13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*

14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*

15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*

16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*

17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*

18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*

19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*

20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*

21: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*

22: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*

23: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*

24: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*

25: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*

26: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*

27: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1929	100.0	347	22	US-09-890-752A-1
2	1879	97.4	389	21	US-09-718-095-18
3	1870	96.9	389	3	US-07-847-951A-216
4	1870	96.9	389	3	US-07-847-951A-219
5	1870	96.9	389	4	US-08-036-218-214
6	1870	96.9	389	4	US-08-036-218-217

7	1868	96.8	389	7	US-08-360-107-116	Sequence 116, App
8	1868	96.8	389	8	US-08-470-896-106	Sequence 106, App
9	1868	96.8	389	8	US-08-471-913-106	Sequence 106, App
10	1868	96.8	389	8	US-08-475-668-106	Sequence 106, App
11	1868	96.8	389	8	US-08-484-223-106	Sequence 106, App
12	1868	96.8	389	8	US-08-484-223A-106	Sequence 106, App
13	1868	96.8	389	8	US-08-485-546-106	Sequence 106, App
14	1868	96.8	389	8	US-08-485-546A-106	Sequence 106, App
15	1868	96.8	389	8	US-08-487-266-106	Sequence 106, App
16	1868	96.8	389	8	US-08-487-266A-106	Sequence 106, App
17	1868	96.8	389	8	US-08-487-355-106	Sequence 106, App
18	1868	96.8	389	8	US-08-487-355A-106	Sequence 106, App
19	1868	96.8	389	13	US-08-919-600-106	Sequence 106, App
20	1868	96.8	389	15	US-09-502-445-106	Sequence 106, App
21	1868	96.8	389	26	US-10-267-682-106	Sequence 106, App
22	1868	96.8	389	26	US-10-267-748-106	Sequence 106, App
23	1868	96.8	389	22	US-09-821-877-2	Sequence 2, App1
24	1868	96.7	389	21	US-09-718-095-19	Sequence 2, App1
25	1866	96.7	389	21	US-09-718-095-17	Sequence 2, App1
26	1743	90.5	399	21	US-09-718-095-5	Sequence 3, App1
27	1743	90.4	399	21	US-09-719-528-3	Sequence 3, App1
28	1721	89.2	400	21	US-10-209-264-3	Sequence 3, App1
29	1721	89.2	400	21	US-09-718-095-15	Sequence 3, App1
30	1721	89.2	400	21	US-09-718-095-16	Sequence 3, App1
31	1721	89.2	400	21	US-09-718-095-20	Sequence 15, App1
32	1721	89.2	400	21	US-08-036-218-212	Sequence 20, App1
33	1700	88.1	400	16	US-09-021-789-6	Sequence 214, App
34	1684	87.3	400	21	US-09-724-890-12	Sequence 12, App1
35	1587	82.3	400	21	US-09-724-890-12	Sequence 12, App1
36	1539	79.8	281	3	US-07-847-951A-214	Sequence 12, App1
37	1539	79.8	281	4	US-08-036-218-212	Sequence 12, App1
38	1534	79.5	281	14	US-09-021-789-6	Sequence 12, App1
39	1534	79.5	281	15	US-09-724-890-12	Sequence 12, App1
40	1534	79.5	281	21	US-09-724-890-12	Sequence 12, App1
41	1534	79.5	281	21	US-09-724-890-6	Sequence 12, App1
42	1534	79.5	281	21	US-09-724-890-12	Sequence 12, App1
43	1460	75.7	281	9	US-08-537-814A-35	Sequence 35, App1
44	1460	75.7	281	13	US-08-942-489-35	Sequence 35, App1
45	1460	75.7	281	15	US-09-182-186A-35	Sequence 35, App1

ALIGNMENTS

RESULT 1

US-09-890-752A-1

Sequence 1, Application US/09890752A

GENERAL INFORMATION:

APPLICANT: Hildt, Eberhard

APPLICANT: Hofschneider, Peter

TITLE OF INVENTION: Particles for Gene Therapy

FILE REFERENCE: 107070-120 (V06-012)

CURRENT APPLICATION NUMBER: US/09/890,752A

CURRENT FILING DATE: 2001-08-03

PRIOR APPLICATION NUMBER: PCT/DE00/00363

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: DE 199 04 800.2

PRIOR FILING DATE: 1999-02-05

NUMBER OF SEQ ID NOS: 21

SOFTWARE: PatentIn version 3.1

SEQ ID NO 1

LENGTH: 347

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURES:

OTHER INFORMATION: Fusion protein comprising a LHBs and heterologous binding

US-09-890-752A-1

Query Match 100.0%; Score 1929; DB 22; Length 347;

Best Local Similarity 100.0%; Pred. No. 4.3e-141;

Matches 347; Conservative 0; Mismatches 0; Gaps 0;

QY	1	MGRDGAAGATGATGTPPHGGLGMSPOAQGILETLPANPPASTNRSGQPTPLSPRL	60
Db	1	MGRDGAAGATGATGTPPHGGLGMSPOAQGILETLPANPPASTNRSGQPTPLSPRL	60
QY	61	NTHGQAMQNSTTHQTLQDPRVGLYFPAGSSSGTVNPVPTVPSISIFSRIGDPAL	120
Db	61	NTHGQAMQNSTTHQTLQDPRVGLYFPAGSSSGTVNPVPTVPSISIFSRIGDPAL	120
QY	121	NMENTSGFLPPLVLOAGFPLTRILITIPQSLDSWMTSLNFGCTTVCCGQNSQSTSN	180
Db	121	NMENTSGFLPPLVLOAGFPLTRILITIPQSLDSWMTSLNFGCTTVCCGQNSQSTSN	180
QY	181	HSPSCPTCGCYRMCRRPFIIFELILLCILFILLVLDYQMLPVCPLIPGSSSTTSG	240
Db	181	HSPSCPTCGCYRMCRRPFIIFELILLCILFILLVLDYQMLPVCPLIPGSSSTTSG	240
QY	241	PCRTCTTAQGTSMYPSCCCTKPSDGNCTCIPPSNAFGFTMEMSARFSWLSLVFP	300
Db	241	PCRTCTTAQGTSMYPSCCCTKPSDGNCTCIPPSNAFGFTMEMSARFSWLSLVFP	300
QY	301	VQMFVGLSPTVWLSVIMMMWYMGPSLSYSILSPFLDLIPFCLMWYI	347
Db	301	VQMFVGLSPTVWLSVIMMMWYMGPSLSYSILSPFLDLIPFCLMWYI	347
RESULT 2			
US-09-718-095-18			
Sequence 18, Application US/09718095			
GENERAL INFORMATION:			
APPLICANT: STUYVER Lleven			
APPLICANT: VAN GEYT Caroline			
APPLICANT: DE GENDT Silja			
TITLE OF INVENTION: New HBV Sequences			
FILE REFERENCE: 2551-52			
CURRENT APPLICATION NUMBER: US/09/718, 095			
CURRENT FILING DATE: 2000-11-22			
PRIOR APPLICATION NUMBER: EP99870252.6			
PRIOR FILING DATE: 1999-12-03			
PRIOR APPLICATION NUMBER: US60/169, 287			
PRIOR FILING DATE: 1999-12-07			
NUMBER OF SEQ ID NOS: 169			
SOFTWARE: PatentIn Ver. 2.1			
SEQ ID NO 18			
LENGTH: 389			
TYPE: PRT			
ORGANISM: Hepatitis B virus			
US-09-718-095-18			
Query Match 97.4%; Score 1879; DB 21; Length 389;			
Best Local Similarity 98.5%; Pred. No. 3, 7e-137;			
Matches 33; Conservative 2; Mismatches 3; Indels 0; Gaps 0.			
QY	6	GAGAFGLGFTPPHGGILGMSPOAQGILETLPANPPASTNRSGROPTPLSPRLNTHPQ	65
Db	48	GAGAFGLGFTPPHGGILGMSPOAQGILETLPANPPASTNRSGROPTPLSPRLNTHPQ	107
QY	66	AMQNSTTHQTLQDPRVGLYFPAGSSSGTVNPVPTVPSISIFSRIGDPALNMENT	125
Db	108	AMQNSTTHQTLQDPRVGLYFPAGSSSGTVNPVPTVPSISIFSRIGDPALNMENT	167
QY	126	TSGFLGPLLVLOAGFPLTRILITIPQSLDSWMTSLNFGCTTVLQGNQSPSNHSPTS	185
Db	168	TSGFLGPLLVLOAGFPLTRILITIPQSLDSWMTSLNFGCTTVLQGNQSPSNHSPTS	227
QY	186	CPPTCPGVRMCLRRPFIIFELILLCILFILLVLDYQMLPVCPLIPGSSSTTSGPRTC	245
Db	228	CPPTCPGVRMCLRRPFIIFELILLCILFILLVLDYQMLPVCPLIPGSSSTTSGPRTC	287
QY	246	TPPAQGTSMYPSCCCTKPSDGNCTCIPPSNAFGFTMEMSARFSWLSLVFPQMFV	305
Db	288	TPPAQGTSMYPSCCCTKPSDGNCTCIPPSNAFGFTMEMSARFSWLSLVFPQMFV	347
QY	306	GLSPTVWLSVIMMMWYMGPSLSYSILSPFLDLIPFCLMWYI	347

[illegible]

RESULT 4

US-07-847-951A-219
Sequence 219, Application US/07847951A
GENERAL INFORMATION:

APPLICANT: Paoletti, Enzo
TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE
TITLE OF INVENTION: STRAIN
NUMBER OF SEQUENCES: 462

CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford
ADDRESSEE: c/o William S. Frommer
STREET: 530 Fifth Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/847,951A
FILING DATE: 19920326
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2400
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-0712
TELEFAX: (212) 840-3333
INFORMATION FOR SEQ ID NO: 219:

SEQUENCE CHARACTERISTICS:
LENGTH: 389 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
US-07-847-951A-219

Query Match Best Local Similarity 96.9%; Score 1870; DB 3; Length 389;

Matches 335; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 6 GAGAFGLGFTPEHGLGWSPOAGILFTLTPANPPASTNRSGRPTPLSPPLNTHQ 65
DB 48 GAGAFGLGFTPEHGLGWSPOAGILFTLTPANPPASTNRSGRPTPLSPPLNTHQ 107
QY 66 AMQWNTTTHQTLQDPRVGLYFPAGSSSGTVNVPPTVASPISIFSRIGDPALMNTI 125
DB 108 AMQWNTTTHQTLQDPRVGLYFPAGSSSGTVNVPPTVASPISIFSRIGDPALMNTI 167
QY 126 TSGFGLPPLVLAQGFLLTRILITIPQSLDSWTSINFLGTTVCCLGONSQSPSHSPTS 185
DB 168 TSGFGLPPLVLAQGFLLTRILITIPQSLDSWTSINFLGTTVCCLGONSQSPSHSPTS 227
QY 186 CPPTCGYRMWCLRRPFIIFLLCLIFLLVLDYQMLPVCPLIPSSSTTSSTGPKCTC 245
DB 228 CPPTCGYRMWCLRRPFIIFLLCLIFLLVLDYQMLPVCPLIPSSSTTSSTGPKCTC 287
QY 246 TTPAGTSMYPSCCCTKSDGNCCTCIPSSMARGKFLMENASARFWSLSLVFVQMFV 305
DB 288 TTPAGTSMYPSCCCTKSDGNCCTCIPSSMARGKFLMENASARFWSLSLVFVQMFV 347
QY 306 GISPFWLVLVIMWYMGPSLSYLSLSPFLPLPLPFCCLMAYI 347
DB 348 GISPFWLVLVIMWYMGPSLSYLSLSPFLPLPLPFCCLMAYI 389

RESULT 5
US-08-036-218-214
Sequence 214, Application US/08036218
GENERAL INFORMATION:

APPLICANT: Paoletti, Enzo
APPLICANT: Perkus, Marion B.
APPLICANT: Taylor, Jill
APPLICANT: Tartaglia, James
APPLICANT: Norton, Elizabeth K.
APPLICANT: Riviere, Michel
APPLICANT: de Taine, Charles
APPLICANT: Limbach, Keith J.
APPLICANT: Johnson, Gerard P.
APPLICANT: Pincus, Steven E.
APPLICANT: Cox, William I.
TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE STRAIN
NUMBER OF SEQUENCES: 221
CORRESPONDENCE ADDRESS:
ADDRESSEE: William S. Frommer c/o
ADDRESSEE: Curtis, Morris & Safford, P.C.
STREET: 530 Fifth Avenue
CITY: New York
STATE: New York
COUNTRY: United State of America
ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/036,218
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/713,967
FILING DATE: 11-JUN-1991
APPLICATION NUMBER: US 07/666,056
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2291
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-0712
TELEFAX: (212) 840-3333
INFORMATION FOR SEQ ID NO: 214:

SEQUENCE CHARACTERISTICS:
LENGTH: 389 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: not relevant
US-08-036-218-214

Query Match Best Local Similarity 96.9%; Score 1870; DB 4; Length 389;

Matches 335; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 6 GAGAFGLGFTPEHGLGWSPOAGILFTLTPANPPASTNRSGRPTPLSPPLNTHQ 65
DB 48 GAGAFGLGFTPEHGLGWSPOAGILFTLTPANPPASTNRSGRPTPLSPPLNTHQ 107
QY 66 AMQWNTTTHQTLQDPRVGLYFPAGSSSGTVNVPPTVASPISIFSRIGDPALMNTI 125
DB 108 AMQWNTTTHQTLQDPRVGLYFPAGSSSGTVNVPPTVASPISIFSRIGDPALMNTI 167
QY 126 TSGFGLPPLVLAQGFLLTRILITIPQSLDSWTSINFLGTTVCCLGONSQSPSHSPTS 185
DB 168 TSGFGLPPLVLAQGFLLTRILITIPQSLDSWTSINFLGTTVCCLGONSQSPSHSPTS 227
QY 186 CPPTCGYRMWCLRRPFIIFLLCLIFLLVLDYQMLPVCPLIPSSSTTSSTGPKCTC 245
DB 228 CPPTCGYRMWCLRRPFIIFLLCLIFLLVLDYQMLPVCPLIPSSSTTSSTGPKCTC 287
QY 246 TTPAGTSMYPSCCCTKSDGNCCTCIPSSMARGKFLMENASARFWSLSLVFVQMFV 305
DB 288 TTPAGTSMYPSCCCTKSDGNCCTCIPSSMARGKFLMENASARFWSLSLVFVQMFV 347

QY 306 GLSPYVLSVIMMMWYGPSTLSLSPFLPLPIFFCLMVYI 347
Db 348 GLSPYVLSVIMMMWYGPSTLSLSPFLPLPIFFCLMVYI 389

RESULT 6
US-08-036-218-217
Sequence 217, Application US/08036218

GENERAL INFORMATION:
APPLICANT: Paolucci, Enzo
APPLICANT: Perkus, Marion E.
APPLICANT: Taylor, Jill
APPLICANT: Taragila, James
APPLICANT: Norton, Elizabeth K.
APPLICANT: Riviere, Michel
APPLICANT: de Talame, Charles
APPLICANT: Limbach, Keith J.
APPLICANT: Johnson, Gerard P.
APPLICANT: Pincus, Steven E.
APPLICANT: Cox, William I.
TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE STRAIN
NUMBER OF SEQUENCES: 221
CORRESPONDENCE ADDRESSES:
ADDRESSEE: William S. Frommer c/o
ADDRESSEE: Curtiss, Morris & Safford, P.C.
STREET: 530 Fifth Avenue
CITY: New York
STATE: New York
COUNTRY: United State of America
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/036,218
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/713,967
FILING DATE: 11-JUN-1991
APPLICATION NUMBER: US 07/666,056
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2291
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 217:
SEQUENCE CHARACTERISTICS:
LENGTH: 389 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: not relevant
US-08-036-218-217

Query Match 96.9%; Score 1870; DB 4; Length 389;
Best Local Similarity 98.0%; Pred. No. 1.8e-136;
Matches 335; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 6 GAGAGFGLFTPPHGLGLGMSPOAGILFTLPANPPASTNRSGROPPLSPPLRNTHQ 65
Db 48 GAGAGFGLFTPPHGLGLGMSPOAGILFTLPANPPASTNRSGROPPLSPPLRNTHQ 107
QY 66 AMQNSTTTHQTLQDPRVGLVFPAGSSSGVTNVPPTVSPISIFSRIGDPALNMENI 125
Db 108 AMQNSTTTHQTLQDPRVGLVFPAGSSSGVTNVPPTVSPISIFSRIGDPALNMENI 167
QY 126 TSGFGLPLLVLAGFFLTRLITIPQSLDSWMTSLNFIAGITVCLGNSQSPTSNHSPTS 185

Db 168 TSGFGLPLLVLAGFFLTRLITIPQSLDSWMTSLNFIAGITVCLGNSQSPTSNHSPTS 227
QY 186 CPPTCGYRMWCLRRPIIFLLICLIFLVLDYQMLVPCPIIPSSSTNGPCRTC 245
Db 228 CPPTCGYRMWCLRRPIIFLLICLIFLVLDYQMLVPCPIIPSSSTNGPCRTC 287
QY 246 TTPAOSTMYPSCCCTKPSDGNCTCIPSSWAFGKFLMENSARFSLVLPVQMFV 305
Db 288 MTTAOSTMYPSCCCTKPSDGNCTCIPSSWAFGKFLMENSARFSLVLPVQMFV 347
QY 306 GLSPYVLSVIMMMWYGPSTLSLSPFLPLPIFFCLMVYI 347
Db 348 GLSPYVLSVIMMMWYGPSTLSLSPFLPLPIFFCLMVYI 389

RESULT 7
US-08-360-107-116
Sequence 116, Application US/08360107

GENERAL INFORMATION:
APPLICANT: Bolognesi, Danl P.
APPLICANT: Matthews, Thomas J.
APPLICANT: Wild, Carl T.
APPLICANT: Barney, Shawn O.
APPLICANT: Lambert, Dennis M.
APPLICANT: Petteway, Stephen R.
APPLICANT: Langlois, Alphonse J.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
TITLE OF INVENTION: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
NUMBER OF SEQUENCES: 137
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/360,107
FILING DATE: 20-DEC-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-013
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
INFORMATION FOR SEQ ID NO: 116:
SEQUENCE CHARACTERISTICS:
LENGTH: 389 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
US-08-360-107-116

Query Match 96.8%; Score 1868; DB 7; Length 389;
Best Local Similarity 98.2%; Pred. No. 2.6e-136;
Matches 336; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 6 GAGAGFGLFTPPHGLGLGMSPOAGILFTLPANPPASTNRSGROPPLSPPLRNTHQ 65
Db 48 GAGAGFGLFTPPHGLGLGMSPOAGILFTLPANPPASTNRSGROPPLSPPLRNTHQ 107
QY 66 AMQNSTTTHQTLQDPRVGLVFPAGSSSGVTNVPPTVSPISIFSRIGDPALNMENI 125

Db 108 AMONSTTHQTLQDPRVGLYFPAGSSSGTVNPVLTASPLSSISFRIGDPALNMENI 167
 QY 126 TSGFLGPLVLVQAGFFLLRLITIPQSLDSWTSINFLGGTIVCLGONSOSPNSHPTS 185
 Db 168 TSGFLGPLVLVQAGFFLLRLITIPQSLDSWTSINFLGGTIVCLGONSOSPNSHPTS 227
 QY 186 CPPTCGRYMNCRLRRFIIFLLCLIFLLVLLDYQMLPVCPLIPGSSSTTSGPCRTC 245
 Db 228 CPPTCGRYMNCRLRRFIIFLLCLIFLLVLLDYQMLPVCPLIPGSSSTTSGPCRTC 287
 QY 246 TTPAGTSMYPSCCCTKPSDGNCTCIPSSWAFKFLMNASARFWSLSLVPVQMFV 305
 Db 288 MTAGTSMYPSCCCTKPSDGNCTCIPSSWAFKFLMNASARFWSLSLVPVQMFV 347
 QY 306 GLSPTVWL SVIMMMYMGPSLSYLSLSPFLPLPIFFCLMVTYI 347
 Db 348 GLSPTVWL SVIMMMYMGPSLSYLSLSPFLPLPIFFCLMVTYI 389

RESULT 8

US-08-470-896-106
 ; Sequence 106, Application US/08470896
 ; GENERAL INFORMATION:
 ; APPLICANT: Bolognesi, Dani P.
 ; APPLICANT: Matthews, Thomas J.
 ; APPLICANT: Wild, Carl T.
 ; APPLICANT: Barney, Shawn O.
 ; APPLICANT: Lambert, Dennis M.
 ; APPLICANT: Peteway, Stephen R.
 ; APPLICANT: Langlois, Alphonse J.
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
 ; TITLE OF INVENTION: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
 ; NUMBER OF INVENTION: TRANSMISSION
 ; NUMBER OF SEQUENCES: 273
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/470,896
 ; FILING DATE: 06-JUN-1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Coruzzi, Laura A.
 ; REGISTRATION NUMBER: 30,742
 ; REFERENCE/DOCKET NUMBER: 7872-020
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 790-9090
 ; TELEFAX: (212) 869-9741/8864
 ; INFORMATION FOR SEQ ID NO: 106:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 389 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: protein
 ; US-08-470-896-106

Query Match 96.8%; Score 1868; DB 8; Length 389;
 Best Local Similarity 98.2%; Pred. No. 2.6e-136;
 Matches 336; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Db 48 GAGAGLGFTHPHGGLGWSPOAGILQTLPANPPASTNRSGRQPTLSPPLRANTHQ 107
 QY 66 AMONSTTHQTLQDPRVGLYFPAGSSSGTVNPVLTASPLSSISFRIGDPALNMENI 125
 Db 108 AMONSTTHQTLQDPRVGLYFPAGSSSGTVNPVLTASPLSSISFRIGDPALNMENI 167
 QY 126 TSGFLGPLVLVQAGFFLLRLITIPQSLDSWTSINFLGGTIVCLGONSOSPNSHPTS 185
 Db 168 TSGFLGPLVLVQAGFFLLRLITIPQSLDSWTSINFLGGTIVCLGONSOSPNSHPTS 227
 QY 186 CPPTCGRYMNCRLRRFIIFLLCLIFLLVLLDYQMLPVCPLIPGSSSTTSGPCRTC 245
 Db 228 CPPTCGRYMNCRLRRFIIFLLCLIFLLVLLDYQMLPVCPLIPGSSSTTSGPCRTC 287
 QY 246 TTPAGTSMYPSCCCTKPSDGNCTCIPSSWAFKFLMNASARFWSLSLVPVQMFV 305
 Db 288 MTAGTSMYPSCCCTKPSDGNCTCIPSSWAFKFLMNASARFWSLSLVPVQMFV 347
 QY 306 GLSPTVWL SVIMMMYMGPSLSYLSLSPFLPLPIFFCLMVTYI 347
 Db 348 GLSPTVWL SVIMMMYMGPSLSYLSLSPFLPLPIFFCLMVTYI 389

RESULT 9

US-08-471-913-106
 ; Sequence 106, Application US/08471913
 ; GENERAL INFORMATION:
 ; APPLICANT: Bolognesi, Dani P.
 ; APPLICANT: Matthews, Thomas J.
 ; APPLICANT: Wild, Carl T.
 ; APPLICANT: Barney, Shawn O.
 ; APPLICANT: Lambert, Dennis M.
 ; APPLICANT: Peteway, Stephen R.
 ; APPLICANT: Langlois, Alphonse J.
 ; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE
 ; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING EBSTEIN-BARR VIRUS
 ; NUMBER OF INVENTION: TRANSMISSION
 ; NUMBER OF SEQUENCES: 209
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/471,913
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Coruzzi, Laura A.
 ; REGISTRATION NUMBER: 30,742
 ; REFERENCE/DOCKET NUMBER: 7872-030
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 790-9090
 ; TELEFAX: (212) 869-9741/8864
 ; INFORMATION FOR SEQ ID NO: 106:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 389 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: protein
 ; US-08-471-913-106

Query Match 96.8%; Score 1868; DB 8; Length 389;

Best Local Similarity 98.2%; Pred. No. 2.6e-136;
Matches 336; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 6 GAGAFGLFTPHHGLGWSPOAGILETLPPANPPASTNROSGOPTPLSPPLRNTHPQ 65
Db 48 GAGAFGLFTPHHGLGWSPOAGILETLPPANPPASTNROSGOPTPLSPPLRNTHPQ 107
QY 66 AMQNSTTFPHQLODPRVGLYFPAGSSSGTVNPPVTVSPISISIFSRIGDPAIEMENI 125
Db 108 AMQNSTTFPHQLODPRVGLYFPAGSSSGTVNPPVTVSPISISIFSRIGDPAIEMENI 167
QY 126 TSGFLGFLVLVQAGFLPLTRILITIPQSLDSWMTSLNFIQTTVCCLQONSQSPISNHSPTS 185
Db 168 TSGFLGFLVLVQAGFLPLTRILITIPQSLDSWMTSLNFIQTTVCCLQONSQSPISNHSPTS 227
QY 186 CPPTCGRYMCLRRPFIIFPLILLCIFLLVLDYQGMIPVCPILPGSSTTSGPCRTC 245
Db 228 CPPTCGRYMCLRRPFIIFPLILLCIFLLVLDYQGMIPVCPILPGSSTTSGPCRTC 287
QY 246 TTPAGTSMYPSCCCTKPSDGNCTCIPISSWAFGKFLMWSARFSWLSLLVPVQMFV 305
Db 288 MTTAGTSMYPSCCCTKPSDGNCTCIPISSWAFGKFLMWSARFSWLSLLVPVQMFV 347
QY 306 GLSPTVWLSVIMMMWYWGPSLYSLSPFLPLPIFCLMWYI 347
Db 348 GLSPTVWLSVIMMMWYWGPSLYSLSPFLPLPIFCLMWYI 389

RESULT 10

US-08-475-668-106
Sequence 106, Application US/08475668

GENERAL INFORMATION:

APPLICANT: Bolognesi, Dani P.
APPLICANT: Matthews, Thomas J.
APPLICANT: Wild, Carl T.
APPLICANT: Barney, Shawn O.
APPLICANT: Lambert, Dennis M.
APPLICANT: Peteway, Stephen R.
APPLICANT: Langlois, Alphonse J.
TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA
TITLE OF INVENTION: VIRUS TRANSMISSION
NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:
ADDRESSER: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,668
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 106:
SEQUENCE CHARACTERISTICS:
LENGTH: 389 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown

MOLECULE TYPE: protein

US-08-475-668-106

Query Match 96.8%; Score 1868; DB 8; Length 389;
Best Local Similarity 98.2%; Pred. No. 2.6e-136;
Matches 336; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 6 GAGAFGLFTPHHGLGWSPOAGILETLPPANPPASTNROSGOPTPLSPPLRNTHPQ 65
Db 48 GAGAFGLFTPHHGLGWSPOAGILETLPPANPPASTNROSGOPTPLSPPLRNTHPQ 107
QY 66 AMQNSTTFPHQLODPRVGLYFPAGSSSGTVNPPVTVSPISISIFSRIGDPAIEMENI 125
Db 108 AMQNSTTFPHQLODPRVGLYFPAGSSSGTVNPPVTVSPISISIFSRIGDPAIEMENI 167
QY 126 TSGFLGFLVLVQAGFLPLTRILITIPQSLDSWMTSLNFIQTTVCCLQONSQSPISNHSPTS 185
Db 168 TSGFLGFLVLVQAGFLPLTRILITIPQSLDSWMTSLNFIQTTVCCLQONSQSPISNHSPTS 227
QY 186 CPPTCGRYMCLRRPFIIFPLILLCIFLLVLDYQGMIPVCPILPGSSTTSGPCRTC 245
Db 228 CPPTCGRYMCLRRPFIIFPLILLCIFLLVLDYQGMIPVCPILPGSSTTSGPCRTC 287
QY 246 TTPAGTSMYPSCCCTKPSDGNCTCIPISSWAFGKFLMWSARFSWLSLLVPVQMFV 305
Db 288 MTTAGTSMYPSCCCTKPSDGNCTCIPISSWAFGKFLMWSARFSWLSLLVPVQMFV 347
QY 306 GLSPTVWLSVIMMMWYWGPSLYSLSPFLPLPIFCLMWYI 347
Db 348 GLSPTVWLSVIMMMWYWGPSLYSLSPFLPLPIFCLMWYI 389

RESULT 11

US-08-484-223-106
Sequence 106, Application US/08484223

GENERAL INFORMATION:

APPLICANT: Bolognesi, Dani P.
APPLICANT: Matthews, Thomas J.
APPLICANT: Wild, Carl T.
APPLICANT: Barney, Shawn O.
APPLICANT: Lambert, Dennis M.
APPLICANT: Peteway, Stephen R.
APPLICANT: Langlois, Alphonse J.
TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
TITLE OF INVENTION: TRANSMISSION
NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:
ADDRESSER: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,223
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-029
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 106:
SEQUENCE CHARACTERISTICS:

LENGTH: 389 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-484-223-106

Query Match 96.8%; Score 1868; DB 8; Length 389;
Best Local Similarity 98.2%; Pred. No. 2, 6e-136;
Matches 336; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 6 GAGAFGLGFTPPHGGILGMSPOAGIILETL PANPPASTNRQSGROPTPLSPPLRNTHPQ 65
DB 48 GAGAFGLGFTPPHGGILGMSPOAGIILETL PANPPASTNRQSGROPTPLSPPLRNTHPQ 107
QY 66 AMQNSTTFHQTLDPRVRLGYFPAGSSSGTVPVPTVSPISISIFSRIGDPALNMENI 125
DB 108 AMQNSTTFHQTLDPRVRLGYFPAGSSSGTVPVPTVSPISISIFSRIGDPALNMENI 167
QY 126 TSGFLGFLVLVQAGFLLTRILTIPOSLSWMTSLNFLAGGTVCLGQNSQSPSNHSPTS 185
DB 168 TSGFLGFLVLVQAGFLLTRILTIPOSLSWMTSLNFLAGGTVCLGQNSQSPSNHSPTS 227
QY 186 CPPTCGYRMCLRRFIFLLCLIFLLVLDYQGMPLVPCPLIPGSSITSGPCRTC 245
DB 228 CPPTCGYRMCLRRFIFLLCLIFLLVLDYQGMPLVPCPLIPGSSITSGPCRTC 287
QY 246 TTPAOGTSMYPSCCCTKPSDGNCTCIPISSWAFGKFLMWASARPSWLSLVPVQMFV 305
DB 288 MTTAOGTSMYPSCCCTKPSDGNCTCIPISSWAFGKFLMWASARPSWLSLVPVQMFV 347
QY 306 GLSPTWMLSVIWMWYMGPSLYSILSPFLPLPIFCLMWYI 347
DB 348 GLSPTWMLSVIWMWYMGPSLYSILSPFLPLPIFCLMWYI 389

RESULT 12

US-08-484-223A-106
Sequence 106, Application US/08484223A
GENERAL INFORMATION:

APPLICANT: Bolognesi, Dant P.
APPLICANT: Matthews, Thomas J.
APPLICANT: Wild, Carl T.
APPLICANT: Barney, Shawn O.
APPLICANT: Lambert, Dennis M.
APPLICANT: Petteway, Stephen R.
APPLICANT: Langlois, Alphonse J.
TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
TRANSMISSION
TITLE OF INVENTION: TRANSMISSION
NUMBER OF SEQUENCES: 239
CORRESPONDENCE ADDRESS:
ADDRESSER: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,223A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-029
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 106:
SEQUENCE CHARACTERISTICS:
LENGTH: 389 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-484-223A-106

Query Match 96.8%; Score 1868; DB 8; Length 389;
Best Local Similarity 98.2%; Pred. No. 2, 6e-136;
Matches 336; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 6 GAGAFGLGFTPPHGGILGMSPOAGIILETL PANPPASTNRQSGROPTPLSPPLRNTHPQ 65
DB 48 GAGAFGLGFTPPHGGILGMSPOAGIILETL PANPPASTNRQSGROPTPLSPPLRNTHPQ 107
QY 66 AMQNSTTFHQTLDPRVRLGYFPAGSSSGTVPVPTVSPISISIFSRIGDPALNMENI 125
DB 108 AMQNSTTFHQTLDPRVRLGYFPAGSSSGTVPVPTVSPISISIFSRIGDPALNMENI 167
QY 126 TSGFLGFLVLVQAGFLLTRILTIPOSLSWMTSLNFLAGGTVCLGQNSQSPSNHSPTS 185
DB 168 TSGFLGFLVLVQAGFLLTRILTIPOSLSWMTSLNFLAGGTVCLGQNSQSPSNHSPTS 227
QY 186 CPPTCGYRMCLRRFIFLLCLIFLLVLDYQGMPLVPCPLIPGSSITSGPCRTC 245
DB 228 CPPTCGYRMCLRRFIFLLCLIFLLVLDYQGMPLVPCPLIPGSSITSGPCRTC 287
QY 246 TTPAOGTSMYPSCCCTKPSDGNCTCIPISSWAFGKFLMWASARPSWLSLVPVQMFV 305
DB 288 MTTAOGTSMYPSCCCTKPSDGNCTCIPISSWAFGKFLMWASARPSWLSLVPVQMFV 347
QY 306 GLSPTWMLSVIWMWYMGPSLYSILSPFLPLPIFCLMWYI 347
DB 348 GLSPTWMLSVIWMWYMGPSLYSILSPFLPLPIFCLMWYI 389

RESULT 13

US-08-485-546-106
Sequence 106, Application US/08485546
GENERAL INFORMATION:

APPLICANT: Bolognesi, Dant P.
APPLICANT: Matthews, Thomas J.
APPLICANT: Wild, Carl T.
APPLICANT: Barney, Shawn O.
APPLICANT: Lambert, Dennis M.
APPLICANT: Petteway, Stephen R.
APPLICANT: Langlois, Alphonse J.
TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE
FUSION-ASSOCIATED EVENTS, INCLUDING EPSTEIN-BARR VIRUS
TRANSMISSION
TITLE OF INVENTION: TRANSMISSION
NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:
ADDRESSER: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,546
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.

```

;
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELETYPE: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 106:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 389 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-485-546-106

Query Match          96.8%; Score 1868; DB 8; Length 389;
Best Local Similarity 98.2%; Pred. No. 2.6e-136;
Matches 336; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 6 GAGAFGLGFTPPHGGILGWSPOAQGILETLPANPPASTNRQSGRQPTPLSPRLNTHQ 65
DB 48 GAGAFGLGFTPPHGGILGWSPOAQGILQTLPANPPASTNRQSGRQPTPLSPRLNTHQ 107
QY 66 AMQNSTTFHQTLODPRVRLGYPAGSSSGTVNPVTTVSPISSIFSRIGDPAIEMNI 125
DB 108 AMQNSTTFHQTLODPRVRLGYPAGSSSGTVNPVTTVSPISSIFSRIGDPAIEMNI 167
QY 126 TSGFLGPLVLVQAGFFLTLRLITIPQSLDSWMTSLNFGATTVCIGNSQSPHSNPS 185
DB 168 TSGFLGPLVLVQAGFFLTLRLITIPQSLDSWMTSLNFGATTVCIGNSQSPHSNPS 227
QY 186 CPPTCGYRMCLRRFIIFLLILCLIFLLVLDYQGMIPVCPLIPGSSITGPCRTC 245
DB 228 CPPTCGYRMCLRRFIIFLLILCLIFLLVLDYQGMIPVCPLIPGSSITGPCRTC 287
QY 246 TTPAOGTSMVPSCCCTKPSDNCCTCIPSSMAFGKFLMEMASARFWSLSLVPFQW 305
DB 288 MTTAOGTSMVPSCCCTKPSDNCCTCIPSSMAFGKFLMEMASARFWSLSLVPFQW 347
QY 306 GLSPTVWLSVIMMMWYMGPSLYSILSPFLPLPIFCLWYI 347
DB 348 GLSPTVWLSVIMMMWYMGPSLYSILSPFLPLPIFCLWYI 389

RESULT 14
US-08-485-546-106
; Sequence 106, Application US/08485546A
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Peteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING RESTEIN-BARR VIRUS
; NUMBER OF SEQUENCES: 214
; CORRESPONDENCE ADDRESS:
; ADDRESS: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485.546A
```

```

;
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELETYPE: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 106:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 389 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-485-546A-106

Query Match          96.8%; Score 1868; DB 8; Length 389;
Best Local Similarity 98.2%; Pred. No. 2.6e-136;
Matches 336; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 6 GAGAFGLGFTPPHGGILGWSPOAQGILETLPANPPASTNRQSGRQPTPLSPRLNTHQ 65
DB 48 GAGAFGLGFTPPHGGILGWSPOAQGILQTLPANPPASTNRQSGRQPTPLSPRLNTHQ 107
QY 66 AMQNSTTFHQTLODPRVRLGYPAGSSSGTVNPVTTVSPISSIFSRIGDPAIEMNI 125
DB 108 AMQNSTTFHQTLODPRVRLGYPAGSSSGTVNPVTTVSPISSIFSRIGDPAIEMNI 167
QY 126 TSGFLGPLVLVQAGFFLTLRLITIPQSLDSWMTSLNFGATTVCIGNSQSPHSNPS 185
DB 168 TSGFLGPLVLVQAGFFLTLRLITIPQSLDSWMTSLNFGATTVCIGNSQSPHSNPS 227
QY 186 CPPTCGYRMCLRRFIIFLLILCLIFLLVLDYQGMIPVCPLIPGSSITGPCRTC 245
DB 228 CPPTCGYRMCLRRFIIFLLILCLIFLLVLDYQGMIPVCPLIPGSSITGPCRTC 287
QY 246 TTPAOGTSMVPSCCCTKPSDNCCTCIPSSMAFGKFLMEMASARFWSLSLVPFQW 305
DB 288 MTTAOGTSMVPSCCCTKPSDNCCTCIPSSMAFGKFLMEMASARFWSLSLVPFQW 347
QY 306 GLSPTVWLSVIMMMWYMGPSLYSILSPFLPLPIFCLWYI 347
DB 348 GLSPTVWLSVIMMMWYMGPSLYSILSPFLPLPIFCLWYI 389

RESULT 15
US-08-485-551-106
; Sequence 106, Application US/08485551
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Peteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA VIRUS
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESS: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```


OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,551
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-023
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-8090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIR
INFORMATION FOR SEQ ID NO: 106:
SEQUENCE CHARACTERISTICS:
LENGTH: 389 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-485-551-106

Query Match 96.8%; Score 1868; DB 8; Length 389;
Best Local Similarity 98.2%; Pred. No. 2.6e-136;
Matches 336; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 6 GAGAFGLGFTPPHGLGWSPOAGILETLPANPPASTNRQSGROPTPLSPPLRYTHPQ 65
DB 48 GAGAFGLGFTPPHGLGWSPOAGILETLPANPPASTNRQSGROPTPLSPPLRYTHPQ 107
QY 66 AMQMNSTTHQTLQDPRVGLFPAGSSSGTVNPVPTVSPISISFRIQDPALMNI 125
DB 108 AMQMNSTTHQTLQDPRVGLFPAGSSSGTVNPVPTVSPISISFRIQDPALMNI 167
QY 126 TSGFLGPLVLQAGFPLTRILTIPOSLSWMTSLNEIGTTVCUGONSQSPSNHSPTS 185
DB 166 TSGFLGPLVLQAGFPLTRILTIPOSLSWMTSLNEIGTTVCUGONSQSPSNHSPTS 227
QY 186 CPPTCGYRMCLRRFIIIFLFIILLCLIFLVLVDYQMLPVCPILPGSSTSTGPCRTC 245
DB 228 CPPTCGYRMCLRRFIIIFLFIILLCLIFLVLVDYQMLPVCPILPGSSTSTGPCRTC 287
QY 246 TTPAOGTSMYPSCCCTKPSDGNCTCIPRSSWAFGKFLMEASARPSWLSLVPVQWFEV 305
DB 288 TTPAOGTSMYPSCCCTKPSDGNCTCIPRSSWAFGKFLMEASARPSWLSLVPVQWFEV 347
QY 306 GLSPTVWLIVIMMMWYWGPSLSYLSLSPFLPLPIFFCLMVIYI 347
DB 348 GLSPTVWLIVIMMMWYWGPSLSYLSLSPFLPLPIFFCLMVIYI 389

Search completed: February 3, 2003, 09:34:07
Job time : 250.857 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 09:25:53 ; Search time 12.6951 Seconds

(without alignments)
551.547 Million cell updates/sec

Title: US-09-890-752A-1

Sequence: 1 MRRGDAAGAFGLGFTPRHGS.....SLSPLPLPLPFLCWMYI 347

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 122226 seqs, 20178551 residues

Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1866	96.7	389	9	US-09-821-877-2
2	1721	89.2	400	9	US-10-209-264-3
3	1534	79.5	281	9	US-09-247-890-12
4	1437	74.5	281	9	US-09-247-890-10
5	1259	65.3	226	9	US-10-104-966-10
6	1259	65.3	226	10	US-09-929-955-10
7	1259	65.3	226	9	US-09-821-877-8
8	1252	64.9	226	10	US-09-812-862-14
9	1231	63.8	397	10	US-09-812-862-6
10	997	51.7	351	10	US-09-812-862-4
11	834.5	43.3	282	9	US-09-247-880-16
12	652	33.8	289	10	US-09-812-862-8
13	649	33.6	346	10	US-09-812-862-2
14	613	31.8	174	10	US-09-818-066-55
15	604	31.3	174	10	US-09-818-066-54
16	579	30.0	174	10	US-09-818-066-56
17	571	29.6	174	10	US-09-818-066-57
18	565	29.3	174	10	US-09-818-066-46
19	561	29.1	174	10	US-09-818-066-48

20	554	28.7	174	10	US-09-818-066-53	Sequence 53, Appl
21	552	28.6	174	10	US-09-818-066-45	Sequence 45, Appl
22	552	28.6	174	10	US-09-818-066-62	Sequence 62, Appl
23	551	28.6	174	10	US-09-818-066-42	Sequence 42, Appl
24	550	28.5	174	10	US-09-818-066-59	Sequence 59, Appl
25	545	28.3	174	10	US-09-818-066-43	Sequence 43, Appl
26	544	28.2	174	10	US-09-818-066-47	Sequence 47, Appl
27	541	28.0	174	10	US-09-818-066-49	Sequence 49, Appl
28	539	27.9	174	10	US-09-818-066-49	Sequence 49, Appl
29	539	27.9	174	10	US-09-818-066-44	Sequence 44, Appl
30	538	27.9	174	10	US-09-818-066-52	Sequence 52, Appl
31	533	27.6	174	10	US-09-818-066-58	Sequence 58, Appl
32	529	27.4	174	10	US-09-818-066-38	Sequence 38, Appl
33	528	27.4	174	10	US-09-818-066-50	Sequence 50, Appl
34	528	27.4	174	10	US-09-818-066-51	Sequence 51, Appl
35	525	27.2	174	10	US-09-818-066-40	Sequence 40, Appl
36	525	27.2	174	10	US-09-818-066-41	Sequence 41, Appl
37	522	27.1	174	10	US-09-818-066-35	Sequence 35, Appl
38	496	25.7	174	10	US-09-818-066-36	Sequence 36, Appl
39	488	25.3	174	10	US-09-818-066-60	Sequence 60, Appl
40	480	24.9	174	10	US-09-818-066-61	Sequence 61, Appl
41	472	24.5	174	10	US-09-818-066-61	Sequence 61, Appl
42	353	18.3	61	9	US-09-821-877-5	Sequence 5, Appl
43	228	11.8	55	10	US-09-879-257A-45	Sequence 45, Appl
44	209	10.8	48	10	US-09-818-066-67	Sequence 67, Appl
45	208.5	10.8	420	10	US-09-812-862-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-09-821-877-2
; Sequence 2, Application US/09821877
; Patent No. US20020177124A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Coleman, Paul F.
; TITLE OF INVENTION: Hepatitis B Virus Surface Antigen Mutant
; TITLE OF INVENTION: And Methods Of Detection Thereof
; FILE REFERENCE: 6794 US-01
; CURRENT APPLICATION NUMBER: US/09/821,877
; CURRENT FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Hepatitis B Virus
US-09-821-877-2

Query Match 96.7%; Score 1866; DB 9; Length 389;
Best Local Similarity 98.2%; Pred. No. 1.9e-132;
Matches 336; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 6 GAGAFGLGFTPRHGSPOAGILETLPANPPASTNRSGRQPTPLSPRLRTHQ 65
DB 48 GAGAFGLGFTPRHGSPOAGILETLPANPPASTNRSGRQPTPLSPRLRTHQ 107
QY 66 AMQWNTTFHQTODPRVRLGLYPAGSSSGTVNPPVTSPISSIFSRIGDPALMNI 125
DB 108 AMQWNTTFHQTODPRVRLGLYPAGSSSGTVNPPVTSPISSIFSRIGDPALMNI 167
QY 126 TSGFLGLVLQAGFLPLRLITIPSLDSWMTSLNLFAGTTCVCGQNSQSPSHSPS 185
DB 168 TSGFLGLVLQAGFLPLRLITIPSLDSWMTSLNLFAGTTCVCGQNSQSPSHSPS 227
QY 186 CPPTCCGYRMCRLRPIIFLLILLCLIFLLVLDYQGLPVCPLIPGSSSTSGCRFC 245
DB 228 CPPTCCGYRMCRLRPIIFLLILLCLIFLLVLDYQGLPVCPLIPGSSSTSGCRFC 287
QY 246 TTPAGQTSVYPSCCCKPDSGNTCTPIPSNAFGKFLMEMASAPRSWLSLLVPVQWFEV 305

Db 288 TTPAGTSMYPSCCCTKPSDGNCTCIPSSNAFGKFLMNASARFSLVLPVQMFV 347
 Qy 306 GLSPTVWLSVIMMMWYMGPSLYSLSPFLPLPIFFCLMWYI 347
 Db 348 GLSPTVWLSVIMMMWYMGPSLYSLSPFLPLPIFFCLMWYI 389

RESULT 2

US-10-209-264-3

Sequence 3, Application US/10209264

Publication No. US2003000311A1

GENERAL INFORMATION:

APPLICANT: Oon, Chong Jin

Ihm, Gek Keow

Zhao, Yi

Chen, Wei Ming

TITLE OF INVENTION: A MUTANT HUMAN HEPATITIS B VIRAL STRAIN AND

USSES THEREOF

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ladas & Parry

STREET: 26 West 61 Street

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10023

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/209,264

FILING DATE: 31-Jul-2002

CLASSIFICATION: 435

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: PCT/SG98/00046

FILING DATE: 19-JAN-1998

ATTORNEY/AGENT INFORMATION:

NAME: Maas, Clifford J.

REGISTRATION NUMBER: 30,086

REFERENCE/DOCKET NUMBER: U-013109-7

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 708-1800

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 400 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-10-209-264-3

Query Match 89.2%; Score 1721; DB 9; Length 400;

Best Local Similarity 90.1%; Pred. No. 1.3e-121;

Matches 308; Conservative 10; Mismatches 24; Indels 0; Gaps 0;

Qy 6 GAGAFGLGFTPPHGLGMSPOAGIILETLRANPPASTNRQSGROPTPLSPRLANTHPQ 65
 Db 59 GVGAFGPGFTPPHGLGMSPOAGIILETLRANPPASTNRQSGROPTPLSPRLANTHPQ 118
 Qy 66 AMONSTTHOTLDDPRVRGLYFPAGSSSGTVNVPPTTSPISSTSRIGDPAIMNENI 125
 Db 119 ATONSTTHOTLDDPRVRGLYFPAGSSSGTVNVPPTTSPISSTSRIGDPAIMNENI 178
 Qy 126 TSGFGLLVLOAGFELTRILITIPQSLDSMTSLNFIAGTTVCLGQNSQSPSTNSHPTS 185
 Db 179 TSGFGLLVLOAGFELTRILITIPQSLDSMTSLNFIAGTTVCLGQNSQSPSTNSHPTS 238
 Qy 186 CPPTCPGYRMWCLRRFIIFLLCLIFLLVLDYQMLPVCPILPGSSTTSGPCRTCT 245
 Db 239 CPPTCPGYRMWCLRRFIIFLLCLIFLLVLDYQMLPVCPILPGSSTTSGPCRTCT 298

Qy 246 TTPAGTSMYPSCCCTKPSDGNCTCIPSSNAFGKFLMNASARFSLVLPVQMFV 305
 Db 299 TTPAGTSMYPSCCCTKPSDGNCTCIPSSNAFGKFLMNASARFSLVLPVQMFV 358
 Qy 306 GLSPTVWLSVIMMMWYMGPSLYSLSPFLPLPIFFCLMWYI 347
 Db 359 GLSPTVWLSVIMMMWYMGPSLYSLSPFLPLPIFFCLMWYI 400

RESULT 3

US-09-247-890-12

Sequence 12, Application US/09247890

Publication No. US20020198162A1

GENERAL INFORMATION:

APPLICANT: Punnonen, Juha

APPLICANT: Baas, Steven H.

APPLICANT: Whalen, Robert Gerald

APPLICANT: Howard, Russell

APPLICANT: Stemmer, Willem P.C.

APPLICANT: Maxygen, Inc.

TITLE OF INVENTION: Antigen library Immunization

FILE REFERENCE: 018097-028710US

CURRENT APPLICATION NUMBER: US/09/247,890

CURRENT FILING DATE: 1999-02-10

EARLIER APPLICATION NUMBER: US 60/074,294

EARLIER FILING DATE: 1998-02-11

EARLIER APPLICATION NUMBER: US 60/105,509

EARLIER FILING DATE: 1998-10-23

NUMBER OF SEQ ID NOS: 25

SOFTWARE: Patent Ver. 2.0

SEQ ID NO 12

LENGTH: 281

TYPE: PRT

ORGANISM: Hepatitis B virus

US-09-247-890-12

Query Match 79.5%; Score 1534; DB 9; Length 281;

Best Local Similarity 98.2%; Pred. No. 8.5e-108;

Matches 276; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 67 MOWNSTTHOTLDDPRVRGLYFPAGSSSGTVNVPPTTSPISSTSRIGDPAIMNENI 126
 Db 1 MOWNSTTHOTLDDPRVRGLYFPAGSSSGTVNVPPTTSPISSTSRIGDPAIMNENI 60
 Qy 127 SGFGLLVLOAGFELTRILITIPQSLDSMTSLNFIAGTTVCLGQNSQSPSTNSHPTS 186
 Db 61 SGFGLLVLOAGFELTRILITIPQSLDSMTSLNFIAGTTVCLGQNSQSPSTNSHPTS 120
 Qy 187 PPTCPGYRMWCLRRFIIFLLCLIFLLVLDYQMLPVCPILPGSSTTSGPCRTCT 246
 Db 121 PPTCPGYRMWCLRRFIIFLLCLIFLLVLDYQMLPVCPILPGSSTTSGPCRTCT 180
 Qy 247 TTPAGTSMYPSCCCTKPSDGNCTCIPSSNAFGKFLMNASARFSLVLPVQMFV 306
 Db 181 TTPAGTSMYPSCCCTKPSDGNCTCIPSSNAFGKFLMNASARFSLVLPVQMFV 240
 Qy 307 LSPTVWLSVIMMMWYMGPSLYSLSPFLPLPIFFCLMWYI 347
 Db 241 LSPTVWLSVIMMMWYMGPSLYSLSPFLPLPIFFCLMWYI 281

RESULT 4
 US-09-247-890-10
 Sequence 10, Application US/09247890
 Publication No. US20020198162A1
 GENERAL INFORMATION:
 APPLICANT: Punnonen, Juha
 APPLICANT: Baas, Steven H.
 APPLICANT: Whalen, Robert Gerald
 APPLICANT: Howard, Russell
 APPLICANT: Stemmer, Willem P.C.
 APPLICANT: Maxygen, Inc.

```

; TITLE OF INVENTION: Antigen Library Immunization
; FILE REFERENCE: 018097-028710US
; CURRENT APPLICATION NUMBER: US/09/247, 890
; CURRENT FILING DATE: 1999-02-10
; EARLIER APPLICATION NUMBER: US 60/074,294
; EARLIER FILING DATE: 1998-02-11
; EARLIER APPLICATION NUMBER: US 60/105,509
; EARLIER FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 10
; LENGTH: 281
; TYPE: PRT
; ORGANISM: Hepatitis B virus
US-09-247-890-10
```

```

Query Match      74.5%; Score 1437; DB 9; Length 281;
Best Local Similarity 91.5%; Pred. No. 1.5e-100;
Matches 257; Conservative 6; Mismatches 18; Indels 0; Gaps 0;
```

```

QY 67 MWNSTTHQTLQDPRVGRGIFPAGSSSGTVNPPTTVSPISISFRLGDPALNMENIT 126
   |||||
DB 1 MWNSTTHQTLQDPRVGRGIFPAGSSSGTVNPPTTVSPISISFRLGDPALNMENIT 60
QY 127 SGFGLPLVLVQAGFLLTRILITIPQSLDSWMTSLNFIQGTTCVCGNSQSPSTSNH 186
   |||||
DB 61 SGFGLPLVLVQAGFLLTRILITIPQSLDSWMTSLNFIQGTTCVCGNSQSPSTSNH 120
QY 187 PPTCPGRWMCRLRRFIFLFLILLCIFLLVLVDYQGMVPCPLIPGSSSTTSQPCRTCT 246
   |||||
DB 121 PPTCPGRWMCRLRRFIFLFLILLCIFLLVLVDYQGMVPCPLIPGSSSTTSQPCRTCT 180
QY 247 TPAQGTSMYPSCCCTKPSDGNCTCIPSSMARGLMEWASARFSLSLVPEV 306
   |||||
DB 181 TPAQGTSMYPSCCCTKPSDGNCTCIPSSMARGLMEWASARFSLSLVPEV 240
QY 307 LSPTVLSVIMMMWYGPISLXSLSPFLPLPIFFCLMWTI 347
   |||||
DB 241 LSPTVLSVIMMMWYGPISLXSLSPFLPLPIFFCLMWTI 281
```

```

RESULT 5
US-10-104-966-10
; Sequence 10, Application US/10104966
; Patent No. US20020155124A1
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUSC1
; CURRENT APPLICATION NUMBER: US/10/104,966
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis B virus S antigen (HbsAg) sequence
US-10-104-966-10
```

```

Query Match      65.3%; Score 1259; DB 9; Length 226;
Best Local Similarity 99.1%; Pred. No. 2.4e-87;
Matches 224; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 122 MENTSGFLPLVLVQAGFLLTRILITIPQSLDSWMTSLNFIQGTTCVCGNSQSPSTSNH 181
   |||||
```

```

DB 1 MENTSGFLPLVLVQAGFLLTRILITIPQSLDSWMTSLNFIQGTTCVCGNSQSPSTSNH 60
QY 182 SPTSCPTCPGRWMCRLRRFIFLFLILLCIFLLVLVDYQGMVPCPLIPGSSSTTSQPC 241
   |||||
DB 61 SPTSCPTCPGRWMCRLRRFIFLFLILLCIFLLVLVDYQGMVPCPLIPGSSSTTSQPC 120
QY 242 CRTCTTPAQGTSMYPSCCCTKPSDGNCTCIPSSMARGLMEWASARFSLSLVPEV 301
   |||||
DB 121 CRTCTTPAQGTSMYPSCCCTKPSDGNCTCIPSSMARGLMEWASARFSLSLVPEV 180
QY 302 QMFVGLSPTVLSVIMMMWYGPISLXSLSPFLPLPIFFCLMWTI 347
   |||||
DB 181 QMFVGLSPTVLSVIMMMWYGPISLXSLSPFLPLPIFFCLMWTI 226
```

```

RESULT 6
US-09-929-955-10
; Sequence 10, Application US/09929955
; Patent No. US20020136740A1
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis B virus S antigen (HbsAg) sequence
US-09-929-955-10
```

```

Query Match      65.3%; Score 1259; DB 10; Length 226;
Best Local Similarity 99.1%; Pred. No. 2.4e-87;
Matches 224; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 122 MENTSGFLPLVLVQAGFLLTRILITIPQSLDSWMTSLNFIQGTTCVCGNSQSPSTSNH 181
   |||||
DB 1 MENTSGFLPLVLVQAGFLLTRILITIPQSLDSWMTSLNFIQGTTCVCGNSQSPSTSNH 60
QY 182 SPTSCPTCPGRWMCRLRRFIFLFLILLCIFLLVLVDYQGMVPCPLIPGSSSTTSQPC 241
   |||||
DB 61 SPTSCPTCPGRWMCRLRRFIFLFLILLCIFLLVLVDYQGMVPCPLIPGSSSTTSQPC 120
QY 242 CRTCTTPAQGTSMYPSCCCTKPSDGNCTCIPSSMARGLMEWASARFSLSLVPEV 301
   |||||
DB 121 CRTCTTPAQGTSMYPSCCCTKPSDGNCTCIPSSMARGLMEWASARFSLSLVPEV 180
QY 302 QMFVGLSPTVLSVIMMMWYGPISLXSLSPFLPLPIFFCLMWTI 347
   |||||
DB 181 QMFVGLSPTVLSVIMMMWYGPISLXSLSPFLPLPIFFCLMWTI 226
```

```

RESULT 7
US-09-821-877-8
; Sequence 8, Application US/09821877
; Patent No. US20020177124A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Coleman, Paul F.
; APPLICANT: Mushahwar, Isha K.
; TITLE OF INVENTION: Hepatitis B Virus Surface Antigen Mutant
US-09-821-877-8
```

Thu Feb 6 14:44:19 2003

us-09-890-752a-1.rapb

Page 4

```

; TITLE OF INVENTION: And Methods Of Detection Thereof
; FILE REFERENCE: 6794.US.01
; CURRENT APPLICATION NUMBER: US/09/821.877
; CURRENT FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PasteSeq for Windows Version 4.0
; SEQ ID NO: 8
; LENGTH: 229
; TYPE: PR1
; ORGANISM: Hepatitis B Virus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (126)...(126)
; OTHER INFORMATION: Xaa = A or T at position 126
; NAME/KEY: VARIANT
; LOCATION: (202)...(202)
; OTHER INFORMATION: Xaa = L or W at position 202
; NAME/KEY: VARIANT
; LOCATION: (210)...(210)
; OTHER INFORMATION: Xaa = T or S at position 210
; US-09-821-877-8

Query Match          65.3%; Score 1259; DB 9; Length 229;
Best Local Similarity 98.3%; Pred. No. 2.5e-87;
Matches 225; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 119 ALNMENTSGFLPLVLDAGFLLRLITIPQSLDSWWTSLNPLGTTVCIGNSQSP 178
DB 1 ARMNENTSGFLPLVLDAGFLLRLITIPQSLDSWWTSLNPLGTTVCIGNSQSP 60
QY 179 SNHSPSCPTCGYRMCRLRPIIFLLCLIFLVLDYQMLPVCPLIPSSSTTS 238
DB 61 SNHSPSCPTCGYRMCRLRPIIFLLCLIFLVLDYQMLPVCPLIPSSSTTS 120
QY 239 TGCRCCTTTPAOGTSMYPSCCCTKPSDGNCTCIPSSNAFGKFLWEMASAFSWLSLLV 298
DB 121 TGCRCCTTTPAOGTSMYPSCCCTKPSDGNCTCIPSSNAFGKFLWEMASAFSWLSLLV 180
QY 299 PFWQFVGLSPTVWLSVYMMWYMGPSLYSILSPFLPLPIFCLWYI 347
DB 181 PFWQFVGLSPTVWLSVYMMWYMGPSLYSILSPFLPLPIFCLWYI 229

RESULT 8
; US-09-812-862-14
; Sequence 14, Application US/09812862
; Patent No. US20020035081A1
; GENERAL INFORMATION:
; APPLICANT: Wanda, Jack R.
; Melegari, Margherita
; TITLE OF INVENTION: INHIBITION OF HEPATITIS B REPLICATION
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/812,862
; FILING DATE: 20-Mar-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/667,073
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 08/492,489
```

```

; FILING DATE: 20-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00786/282001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 226 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 14:
; US-09-812-862-14

Query Match          64.9%; Score 1252; DB 10; Length 226;
Best Local Similarity 98.7%; Pred. No. 8.1e-87;
Matches 223; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 122 MENITSGFLPLVLDAGFLLRLITIPQSLDSWWTSLNPLGTTVCIGNSQSP 181
DB 1 MENITSGFLPLVLDAGFLLRLITIPQSLDSWWTSLNPLGTTVCIGNSQSP 60
QY 182 SPSCPTCGYRMCRLRPIIFLLCLIFLVLDYQMLPVCPLIPSSSTTS 241
DB 61 SPSCPTCGYRMCRLRPIIFLLCLIFLVLDYQMLPVCPLIPSSSTTS 120
QY 242 CRCTTTPAOGTSMYPSCCCTKPSDGNCTCIPSSNAFGKFLWEMASAFSWLSLLV 301
DB 121 CRCTTTPAOGTSMYPSCCCTKPSDGNCTCIPSSNAFGKFLWEMASAFSWLSLLV 180
QY 302 QFWVGLSPTVWLSVYMMWYMGPSLYSILSPFLPLPIFCLWYI 347
DB 181 QFWVGLSPTVWLSVYMMWYMGPSLYSILSPFLPLPIFCLWYI 226

RESULT 9
; US-09-812-862-6
; Sequence 6, Application US/09812862
; Patent No. US20020035081A1
; GENERAL INFORMATION:
; APPLICANT: Wanda, Jack R.
; Melegari, Margherita
; TITLE OF INVENTION: INHIBITION OF HEPATITIS B REPLICATION
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/812,862
; FILING DATE: 20-Mar-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/667,073
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 08/492,489
; FILING DATE: 20-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
```

REFERENCE/DOCKET NUMBER: 00786/282001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 397 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-812-862-6

Query Match
Best Local Similarity 63.8%; Score 1231; DB 10; Length 397;
Matches 237; Conservative 8; Mismatches 31; Indels 64; Gaps 5;

Db 9 AFGIAGF-TPEHGGILGMSPOAGILETLFANPPASTNQSOPPTPLSPPLANTHPOAM 67
121 SFGWIRTP-----AYRPNAPILSTL-----DETIVRRRGSSPRRRTPSPRRRSQS- 170
QY 68 QMNSTTFHOTLQDPVRGILYFPAGSSSGTVNFPPTVSPISIFARIGDPALNMENITS 127
Db 171 -----PRRR-----SQ 177

QY 128 GFLGPLVLAQGFPLTLRLTIPQSLDSWMTSLNPLGTTVCIGONSQSPTNHSPSC 187
Db 178 SRIGPLVLAQGFPLTLRLTIPQSLDSWMTSLNPLGTTVCIGONSQSPTNHSPSC 237

QY 188 PTCGPRWMCRLRFFILFILLCLIFLVLDYOGMLPVCPILGSSSTTSRGPRTCTT 247
Db 238 PTCGPRWMCRLRFFILFILLCLIFLVLDYOGMLPVCPILGSSSTTSRGPRTCTT 297

QY 248 PAQGSMTSPSCCTKPSDGNCTCIPSPSSMAFGKFLMEMASARFSLVLFVQMFVGL 307
Db 298 TPAQGSMTSPSCCTKPSDGNCTCIPSPSSMAFGKFLMEMASARFSLVLFVQMFVGL 357

QY 308 SPTVWLSVIMMMWYMGPSLYSLSPFLPLIFFLCLWYI 347
Db 358 SPTVWLSVIMMMWYMGPSLYSLSPFLPLIFFLCLWYI 397

RESULT 10
US-09-812-862-4
Sequence 4, Application US/09812862
Patent No. US20020035081A1
GENERAL INFORMATION:
APPLICANT: Wanda, Jack R.
Scaglioni, Pier Paolo
Melegari, Margherita
TITLE OF INVENTION: INHIBITION OF HEPATITIS B REPLICATION
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/812,862
FILING DATE: 20-Mar-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/667,073
FILING DATE: <Unknown>
APPLICATION NUMBER: 08/492,489

FILING DATE: 20-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/282001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 351 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-812-862-4

Query Match
Best Local Similarity 51.7%; Score 997; DB 10; Length 351;
Matches 174; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 172 QNSQSPSNHSPSCPPCPGPRWMCRLRFFILFILLCLIFLVLDYOGMLPVCPIL 231
Db 176 QNSQSPSNHSPSCPPCPGPRWMCRLRFFILFILLCLIFLVLDYOGMLPVCPIL 235

QY 232 PGSSSTTSRGPRTCTTPAOGTSMTPSCCTKPSDGNCTCIPSPSSMAFGKFLMEMASARF 291
Db 236 PGSSSTTSRGPRTCTTPAOGTSMTPSCCTKPSDGNCTCIPSPSSMAFGKFLMEMASARF 295

QY 292 SMLSULVLFVQMFVGLSPTVWLSVIMMMWYMGPSLYSLSPFLPLIFFLCLWYI 347
Db 296 SMLSULVLFVQMFVGLSPTVWLSVIMMMWYMGPSLYSLSPFLPLIFFLCLWYI 351

RESULT 11
US-09-247-890-16
Sequence 16, Application US/09247890
Publication No. US20020198162A1
GENERAL INFORMATION:
APPLICANT: Punomen, Usha
APPLICANT: Baas, Steven H.
APPLICANT: Whalen, Robert Gerald
APPLICANT: Howard, Russell
APPLICANT: Stemmer, Willem P.C.
APPLICANT: Maxygen, Inc.
TITLE OF INVENTION: Antigen Library Immunization
FILE REFERENCE: 018097-028710US
CURRENT APPLICATION NUMBER: US/09/247,890
CURRENT FILING DATE: 1999-02-10
EARLIER APPLICATION NUMBER: US 60/074,294
EARLIER FILING DATE: 1998-02-11
EARLIER APPLICATION NUMBER: US 60/105,509
EARLIER FILING DATE: 1998-10-23
NUMBER OF SEQ ID NOS: 25
SOFTWARE: Patent Ver. 2.0
SEQ ID NO: 16
LENGTH: 282
TYPE: PRT
ORGANISM: Woodchuck hepatitis B virus
US-09-247-890-16

Query Match
Best Local Similarity 43.3%; Score 834.5; DB 9; Length 282;
Matches 159; Conservative 29; Mismatches 85; Indels 7; Gaps 2;

QY 73 TRH-----OTLQDPVRGILYFPAGSSSGTVNFPPTVSPISIFSRIGDPALNMENITS 127
Db 5 TRHGFVVDGLHDLTTTERQHNAYGDPFTLSPAVPYSTLSPSTTGDRPLSLEMSPS 64

QY 128 GFLGPLVLAQGFPLTLRLTIPQSLDSWMTSLNPLGTTVCIGONSQSPTNHSPSC 187
Db 65 SLGGLIAGLQVYFLMTYILTIATQNLDMWMTSLSPGSIPECTGONSQFOGCKHLPTSCP 124

188 PTCGVRMCLRRFIFLFLICLIFLVLDYQGMPLVPCPLRSGSTSTGCRCTT 247
125 PTCGFRMWRRIITVILVLLCLIFLVLDYQGMPLVPCPLRSGSTSTGCRCTT 182
248 PAQGTSMYPSCCCTKRPDGNCTCIPSSWAFGKFLMWSARFSLVFPVQFVGL 307
183 SAQMYTPPYCCCKPAGNCTCIPSSWAFGKFLMWSARFSLVFPVQFVGL 242
308 SPTWLVSYIMMMWYSGPSLYSLSPFLPLPIFCLWYI 347
243 SLIAPFLIMIMWFGPALSLIPFPIFVLFLWYI 282

RESULT 12
US-09-812-862-8
Sequence 8, Application US/09812862
Patent No. US20020035081A1
GENERAL INFORMATION:
APPLICANT: Wanda, Jack R.
Scaglioni, Pier Paolo
Melegari, Margherita
TITLE OF INVENTION: INHIBITION OF HEPATITIS B REPLICATION
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/812,862
FILING DATE: 20-Mar-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/667,073
FILING DATE: <Unknown>
APPLICATION NUMBER: 08/492,489
FILING DATE: 20-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/282001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 289 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-812-862-8

Query Match 33.8%; Score 652; DB 10; Length 289;
Best Local Similarity 98.2%; Pred. No. 7.5e-42;
Matches 112; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

234 SSTSTGCRCTTTPAQTSMYPSCCCTKRPDGNCTCIPSSWAFGKFLMWSARFSLV 293
176 SSTSTGCRCTTTPAQTSMYPSCCCTKRPDGNCTCIPSSWAFGKFLMWSARFSLV 235
294 LSLVLPVQWFGVLSPTWLVSYIMMMWYSGPSLYSLSPFLPLPIFCLWYI 347
236 LSLVLPVQWFGVLSPTWLVSYIMMMWYSGPSLYSLSPFLPLPIFCLWYI 289

RESULT 13
US-09-812-862-2
Sequence 2, Application US/09812862
Patent No. US20020035081A1
GENERAL INFORMATION:
APPLICANT: Wanda, Jack R.
Scaglioni, Pier Paolo
Melegari, Margherita
TITLE OF INVENTION: INHIBITION OF HEPATITIS B REPLICATION
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/812,862
FILING DATE: 20-Mar-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/667,073
FILING DATE: <Unknown>
APPLICATION NUMBER: 08/492,489
FILING DATE: 20-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/282001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 346 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-812-862-2

Query Match 33.6%; Score 649; DB 10; Length 346;
Best Local Similarity 63.1%; Pred. No. 1.5e-41;
Matches 111; Conservative 21; Mismatches 42; Indels 2; Gaps 1;

172 QNSQSPNSHSPSCPTGVRMCLRRFIFLFLICLIFLVLDYQGMPLVPCPLI 231
173 QNSQSPNSHSPSCPTGVRMCLRRFIFLFLICLIFLVLDYQGMPLVPCPLI 232
232 PGSSTSTGCRCTTTPAQTSMYPSCCCTKRPDGNCTCIPSSWAFGKFLMWSARF 291
233 PTEETVYV--CROCTISADQMYTPPYCCCKPAGNCTCIPSSWAFGKFLMWSARF 290
292 SMLVLPVQWFGVLSPTWLVSYIMMMWYSGPSLYSLSPFLPLPIFCLWYI 347
291 SMLVLPVQWFGVLSPTWLVSYIMMMWYSGPSLYSLSPFLPLPIFCLWYI 346

RESULT 14
US-09-818-066-55
Sequence 55, Application US/09818066
Patent No. US20020032307A1
GENERAL INFORMATION:
APPLICANT: Shuping Tong et al.

TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR
NUMBER OF SEQUENCES: 75
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/818,066
FILING DATE: 27-Mar-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/683,262
FILING DATE: 18-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 31,819
REFERENCE/DOCKET NUMBER: 00786/287002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 174 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 55:
US-09-818-066-55
Query Match 31.8%; Score 613; DB 10; Length 174;
Best Local Similarity 96.6%; Pred No. 3.5e-39;
Matches 112; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 6 GAGAGLGFTPPHGLGWSPOAGIILTLTPANPPASTNRQSGROPTPLSPPLRNTHPQ 65
DB 59 GAGAGLGFTPPHGLGWSPOAGIILTLTPANPPASTNRQSGROPTPLSPPLRNTHPQ 118
QY 66 AMQNSTTFHOTLQDPRVRGLYFPAGSSSGTVNPVPTTSPISISIFSRIGDPALN 121
DB 119 AMQNSTTFHOTLQDPRVRGLYFPAGSSSGTVNPVPTTSPISISIFSRIGDPALN 174
RESULT 15
US-09-818-066-54
Sequence 54, Application US/09818066
Patent No. US20020032307A1
GENERAL INFORMATION:
APPLICANT: Shuping Tong et al.
TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR
NUMBER OF SEQUENCES: 75
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/818,066

FILING DATE: 27-Mar-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/683,262
FILING DATE: 18-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 31,819
REFERENCE/DOCKET NUMBER: 00786/287002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 174 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 54:
US-09-818-066-54

Query Match 31.3%; Score 604; DB 10; Length 174;
Best Local Similarity 96.6%; Pred. No. 1.7e-38;
Matches 112; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 6 GAGAGLGFTPPHGLGWSPOAGIILTLTPANPPASTNRQSGROPTPLSPPLRNTHPQ 65
DB 59 GAGAGLGFTPPHGLGWSPOAGIILTLTPANPPASTNRQSGROPTPLSPPLRNTHPQ 118
QY 66 AMQNSTTFHOTLQDPRVRGLYFPAGSSSGTVNPVPTTSPISISIFSRIGDPALN 121
DB 119 AMQNSTTFHOTLQDPRVRGLYFPAGSSSGTVNPVPTTSPISISIFSRIGDPALN 174

Search completed: February 3, 2003, 09:35:27
Job time: 13.6951 secs

Db 48 GAGAGGAGGCTPPHGGLLGMSPOAGGILQTLPANPPASTNRQGTGROPTPLSPPLRNTHPQ 107
QY 66 AMONWSTTHQOTLQDPRVGLYFPAGSSSGTVNPPPTVSPISISFSGIDPALNMENI 125
Db 108 AMONWSTTHQOTLQDPRVGLYFPAGSSSGTVNPPPTVSPISISFSGIDPALNMENI 167
QY 126 TSGFLGFLVLVQAGFLPLRLITLITIPQSLDSWMTSLNFIQGTTCVCGQNSQSPTSNHSPTS 185
Db 168 TSGFLGFLVLVQAGFLPLRLITLITIPQSLDSWMTSLNFIQGTTCVCGQNSQSPTSNHSPTS 227
QY 186 CPPTCPGYRMWCLRRPILIFILLICLIFLVLDYQMLPVCPLIPGSSSTSTGBCRTC 245
Db 228 CPPTCPGYRMWCLRRPILIFILLICLIFLVLDYQMLPVCPLIPGSSSTSTGBCRTC 287
QY 246 TTPAOGTSMYPSGCCCTKPSDGNCTCIPISSWAFGKFLMENSARFMSLSLVPFQMFV 305
Db 288 MTTAOGTSMYPSGCCCTKPSDGNCTCIPISSWAFGKFLMENSARFMSLSLVPFQMFV 347
QY 306 GLSPTVWLSVIMMMWYWGPSLYSILSPFLPLPIFCLMWYI 347
Db 348 GLSPTVWLSVIMMMWYWGPSLYSILSPFLPLPIFCLMWYI 389

RESULT 2

US-08-105-483-219
Sequence 219, Application US/08105483
Patent No. 5494807
GENERAL INFORMATION:
APPLICANT: Paolietti, Enzo
TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE
TITLE OF INVENTION: STRAIN
NUMBER OF SEQUENCES: 462
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford
ADDRESSEE: c/o William S. Frommer
STREET: 530 Fifth Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/105,483
FILING DATE: 12-AUG-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/847,951
FILING DATE: 06-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2400
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 219:
SEQUENCE CHARACTERISTICS:
LENGTH: 389 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-105-483-219

Query Match 96.9%; Score 1870; DB 1; Length 389;
Best Local Similarity 98.0%; Pred. No. 3.4e-157;
Matches 335; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
6 GAGAGGAGGCTPPHGGLLGMSPOAGGILQTLPANPPASTNRQGTGROPTPLSPPLRNTHPQ 65

Db 48 GAGAGGAGGCTPPHGGLLGMSPOAGGILQTLPANPPASTNRQGTGROPTPLSPPLRNTHPQ 107
QY 66 AMONWSTTHQOTLQDPRVGLYFPAGSSSGTVNPPPTVSPISISFSGIDPALNMENI 125
Db 108 AMONWSTTHQOTLQDPRVGLYFPAGSSSGTVNPPPTVSPISISFSGIDPALNMENI 167
QY 126 TSGFLGFLVLVQAGFLPLRLITLITIPQSLDSWMTSLNFIQGTTCVCGQNSQSPTSNHSPTS 185
Db 168 TSGFLGFLVLVQAGFLPLRLITLITIPQSLDSWMTSLNFIQGTTCVCGQNSQSPTSNHSPTS 227
QY 186 CPPTCPGYRMWCLRRPILIFILLICLIFLVLDYQMLPVCPLIPGSSSTSTGBCRTC 245
Db 228 CPPTCPGYRMWCLRRPILIFILLICLIFLVLDYQMLPVCPLIPGSSSTSTGBCRTC 287
QY 246 TTPAOGTSMYPSGCCCTKPSDGNCTCIPISSWAFGKFLMENSARFMSLSLVPFQMFV 305
Db 288 MTTAOGTSMYPSGCCCTKPSDGNCTCIPISSWAFGKFLMENSARFMSLSLVPFQMFV 347
QY 306 GLSPTVWLSVIMMMWYWGPSLYSILSPFLPLPIFCLMWYI 347
Db 348 GLSPTVWLSVIMMMWYWGPSLYSILSPFLPLPIFCLMWYI 389

RESULT 3

US-08-709-209-216
Sequence 216, Application US/08709209
Patent No. 5762938
GENERAL INFORMATION:
APPLICANT: Paolietti, Enzo
TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE
TITLE OF INVENTION: STRAIN
NUMBER OF SEQUENCES: 462
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford
ADDRESSEE: c/o William S. Frommer
STREET: 530 Fifth Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/709,209
FILING DATE: 21-AUG-1996
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/105,483
FILING DATE: 12-AUG-1993
APPLICATION NUMBER: US 07/847,951
FILING DATE: 06-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2400
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 216:
SEQUENCE CHARACTERISTICS:
LENGTH: 389 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-709-209-216

Query Match 96.9%; Score 1870; DB 1; Length 389;
Best Local Similarity 98.0%; Pred. No. 3.4e-157;
Matches 335; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 6 GAGAGGAGTTPPHGILGMSPOAGGILFTLPANPPASTNRSGROPTPLSPPLANTHQ 65
DB 48 GAGAGGAGTTPPHGILGMSPOAGGILFTLPANPPASTNRSGROPTPLSPPLANTHQ 107
QY 66 AMONNSTTHQTLQDPRVRLGYPFAGSSSGTVNVPPTVSPISSTRIGDPLANNEMI 125
DB 108 AMONNSTTHQTLQDPRVRLGYPFAGSSSGTVNVPPTVSPISSTRIGDPLANNEMI 167
QY 126 TSGFLGFLVLVQAGFFLLTRILITIPQSLDSWMTSLNFIAGTTVCLGQNSQSPSTNSHPTS 185
DB 168 TSGFLGFLVLVQAGFFLLTRILITIPQSLDSWMTSLNFIAGTTVCLGQNSQSPSTNSHPTS 227
QY 186 CPPTCGYRMWCLRRFIFLFIPLILLCIFILVLDYQMLPVCPLIPGSSSTSTGPCRTC 245
DB 228 CPPTCGYRMWCLRRFIFLFIPLILLCIFILVLDYQMLPVCPLIPGSSSTSTGPCRTC 287
QY 246 TTPAOGTSMYPSGCCCTKPSDNCCTCIPSPSSNARFKFLMNASARFWSLILVFPVQMFV 305
DB 288 MTTAOGTSMYPSGCCCTKPSDNCCTCIPSPSSNARFKFLMNASARFWSLILVFPVQMFV 347
QY 306 GLSPTVWLSVIWMWYMGPSLSYLSLSPFLPLPIFFCLMAYI 347
DB 348 GLSPTVWLSVIWMWYMGPSLSYLSLSPFLPLPIFFCLMAYI 389

RESULT 4
US-08-709-209-219

/ Sequence 219, Application US/08709209
/ Patent No. 5762938
/ GENERAL INFORMATION:
/ APPLICANT: Paoletti, Enzo
/ TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE
/ NUMBER OF SEQUENCES: 462
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Curtis, Morris & Safford
/ ADDRESSEE: c/o William S. Frommer
/ STREET: 530 Fifth Avenue
/ CITY: New York
/ STATE: NY
/ COUNTRY: USA
/ ZIP: 10036
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent in Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/709,209
/ FILING DATE: 21-AUG-1996
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/105,483
/ FILING DATE: 12-AUG-1993
/ APPLICATION NUMBER: US 07/847,951
/ FILING DATE: 06-MAR-1992
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Frommer, William S.
/ REGISTRATION NUMBER: 25,506
/ REFERENCE/DOCKET NUMBER: 454310-2400
/ TELEPHONE: (212) 840-3333
/ TELEFAX: (212) 840-0712
/ INFORMATION FOR SEQ ID NO: 219:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 389 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ US-08-709-209-219

Query Match 96.9%; Score 1870; DB 1, Length 389;
Best Local Similarity 98.0%; Pred. No. 3,4e-157;

Matches 335; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 6 GAGAGGAGTTPPHGILGMSPOAGGILFTLPANPPASTNRSGROPTPLSPPLANTHQ 65
DB 48 GAGAGGAGTTPPHGILGMSPOAGGILFTLPANPPASTNRSGROPTPLSPPLANTHQ 107
QY 66 AMONNSTTHQTLQDPRVRLGYPFAGSSSGTVNVPPTVSPISSTRIGDPLANNEMI 125
DB 108 AMONNSTTHQTLQDPRVRLGYPFAGSSSGTVNVPPTVSPISSTRIGDPLANNEMI 167
QY 126 TSGFLGFLVLVQAGFFLLTRILITIPQSLDSWMTSLNFIAGTTVCLGQNSQSPSTNSHPTS 185
DB 168 TSGFLGFLVLVQAGFFLLTRILITIPQSLDSWMTSLNFIAGTTVCLGQNSQSPSTNSHPTS 227
QY 186 CPPTCGYRMWCLRRFIFLFIPLILLCIFILVLDYQMLPVCPLIPGSSSTSTGPCRTC 245
DB 228 CPPTCGYRMWCLRRFIFLFIPLILLCIFILVLDYQMLPVCPLIPGSSSTSTGPCRTC 287
QY 246 TTPAOGTSMYPSGCCCTKPSDNCCTCIPSPSSNARFKFLMNASARFWSLILVFPVQMFV 305
DB 288 MTTAOGTSMYPSGCCCTKPSDNCCTCIPSPSSNARFKFLMNASARFWSLILVFPVQMFV 347
QY 306 GLSPTVWLSVIWMWYMGPSLSYLSLSPFLPLPIFFCLMAYI 347
DB 348 GLSPTVWLSVIWMWYMGPSLSYLSLSPFLPLPIFFCLMAYI 389

RESULT 5
US-08-458-101-216

/ Sequence 216, Application US/08458101
/ Patent No. 5766599
/ GENERAL INFORMATION:
/ APPLICANT: Paoletti, Enzo
/ APPLICANT: Perkins, Marion E.
/ APPLICANT: Taylor, Jill
/ APPLICANT: Tartaglia, James
/ APPLICANT: No. 5766599ton, Elizabeth K.
/ APPLICANT: Riviere, Michel
/ APPLICANT: de Taise, Charles
/ APPLICANT: Limbach, Keith J.
/ APPLICANT: Johnson, Gerard P.
/ APPLICANT: Pincus, Steven B.
/ APPLICANT: Cox, William I.
/ APPLICANT: Audonnet, Jean-Christophe Francis
/ APPLICANT: Gettig, Russell Robert
/ TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE
/ NUMBER OF SEQUENCES: 467
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Curtis, Morris & Safford
/ ADDRESSEE: c/o William S. Frommer
/ STREET: 530 Fifth Avenue
/ CITY: New York
/ STATE: NY
/ COUNTRY: USA
/ ZIP: 10036
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent in Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/458,101
/ FILING DATE: 01-JUN-1995
/ CLASSIFICATION: 424
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Frommer, William S.
/ REGISTRATION NUMBER: 25,506
/ REFERENCE/DOCKET NUMBER: 454310-2740
/ TELEPHONE: (212) 840-3333
/ TELEFAX: (212) 840-0712
/ INFORMATION FOR SEQ ID NO: 216:

SEQUENCE CHARACTERISTICS:
LENGTH: 389 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-458-101-216

Query Match 96.9%; Score 1870; DB 1; Length 389;
Best Local Similarity 98.0%; Pred. No. 3,4e-157;
Matches 335; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 6 GAGATGCTPPHGGILGNSPOAGILETLPANPPASTNROSGRPTLSPPLANTHQ 65
DB 48 GAGATGCTPPHGGILGNSPOAGILETLPANPPASTNROSGRPTLSPPLANTHQ 107
QY 66 AMQNSTTFPHQTLQDPVRGLYFPAGSSSGTVNPVPTVSPISIFSRIGDPAALMENI 125
DB 108 AMQNSTTFPHQTLQDPVRGLYFPAGSSSGTVNPVPTVSPISIFSRIGDPAALMENI 167
QY 126 TSGFLGLVLYOAGFFLLTRILITIPQSLDSWMTSLNPLGGITVCLGONSQSPSNHSPTS 185
DB 168 TSGFLGLVLYOAGFFLLTRILITIPQSLDSWMTSLNPLGGITVCLGONSQSPSNHSPTS 227
QY 186 CPPTCGRWMCRLRRIFIFILLCLIFLLVLDYQGLPVCPLIPGSSSTSGPCRTC 245
DB 228 CPPTCGRWMCRLRRIFIFILLCLIFLLVLDYQGLPVCPLIPGSSSTSGPCRTC 287
QY 246 TTPAOGTSMYPSCCCTKPSDGNCTCIPSSNAFGKFLWEMASARFWSLLVPVQMFV 305
DB 288 MTTAOGTSMYPSCCCTKPSDGNCTCIPSSNAFGKFLWEMASARFWSLLVPVQMFV 347
QY 306 GLSPTVWLSVIMMMWYGPSLYSLSPFLPLIFPCLMWYI 347
DB 348 GLSPTVWLSVIMMMWYGPSLYSLSPFLPLIFPCLMWYI 389

RESULT 6
US-08-458-101-219
Sequence 219, Application US/08458101

GENERAL INFORMATION:
PATENT: Paoletti, Enzo
APPLICANT: Paoletti, Enzo
APPLICANT: Perkus, Marion E.
APPLICANT: Taylor, Jill
APPLICANT: Tartaglia, James
APPLICANT: No. 5766599ton, Elizabeth K.
APPLICANT: Riviere, Michel
APPLICANT: de Talaue, Charles
APPLICANT: Limbach, Keith J.
APPLICANT: Johnson, Gerard P.
APPLICANT: Pincus, Steven E.
APPLICANT: Cox, William I.
APPLICANT: Audonnet, Jean-Christophe
APPLICANT: Gettlig, Russell Robert
TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE
TITLE OF INVENTION: STRAIN
NUMBER OF SEQUENCES: 467
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtiss, Morris & Safford
ADDRESSEE: c/o William S. Frommer
STREET: 530 Fifth Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,101
FILING DATE: 01-JUN-1995

CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2740
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 219:
SEQUENCE CHARACTERISTICS:
LENGTH: 389 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-458-101-219

Query Match 96.9%; Score 1870; DB 1; Length 389;
Best Local Similarity 98.0%; Pred. No. 3,4e-157;
Matches 335; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 6 GAGATGCTPPHGGILGNSPOAGILETLPANPPASTNROSGRPTLSPPLANTHQ 65
DB 48 GAGATGCTPPHGGILGNSPOAGILETLPANPPASTNROSGRPTLSPPLANTHQ 107
QY 66 AMQNSTTFPHQTLQDPVRGLYFPAGSSSGTVNPVPTVSPISIFSRIGDPAALMENI 125
DB 108 AMQNSTTFPHQTLQDPVRGLYFPAGSSSGTVNPVPTVSPISIFSRIGDPAALMENI 167
QY 126 TSGFLGLVLYOAGFFLLTRILITIPQSLDSWMTSLNPLGGITVCLGONSQSPSNHSPTS 185
DB 168 TSGFLGLVLYOAGFFLLTRILITIPQSLDSWMTSLNPLGGITVCLGONSQSPSNHSPTS 227
QY 186 CPPTCGRWMCRLRRIFIFILLCLIFLLVLDYQGLPVCPLIPGSSSTSGPCRTC 245
DB 228 CPPTCGRWMCRLRRIFIFILLCLIFLLVLDYQGLPVCPLIPGSSSTSGPCRTC 287
QY 246 TTPAOGTSMYPSCCCTKPSDGNCTCIPSSNAFGKFLWEMASARFWSLLVPVQMFV 305
DB 288 MTTAOGTSMYPSCCCTKPSDGNCTCIPSSNAFGKFLWEMASARFWSLLVPVQMFV 347
QY 306 GLSPTVWLSVIMMMWYGPSLYSLSPFLPLIFPCLMWYI 347
DB 348 GLSPTVWLSVIMMMWYGPSLYSLSPFLPLIFPCLMWYI 389

RESULT 7
US-08-486-099-106
Sequence 106, Application US/08486099

GENERAL INFORMATION:
PATENT: Bolognesi, Daniel P.
APPLICANT: Bolognesi, Daniel P.
APPLICANT: Matthews, Thomas J.
APPLICANT: Wild, Carl T.
APPLICANT: Barney, Shawn O.
APPLICANT: Lambert, Dennis M.
APPLICANT: Letteway, Stephen R.
APPLICANT: Langlois, Alphonse J.
TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HEPATITIS
NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,099
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-011
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 106:
SEQUENCE CHARACTERISTICS:
LENGTH: 389 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-486-099-106

Query Match 96.8%; Score 1868; DB 3; Length 389;
Best Local Similarity 98.2%; Pred. No. 5.2e-157;
Matches 336; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 6 GAGAFGLGFTPHHGLGMSPOAGILLETLPANPPASTNRSGROPTPLSPPLRNTHQ 65
DB 48 GAGAFGLGFTPHHGLGMSPOAGILLETLPANPPASTNRSGROPTPLSPPLRNTHQ 107
QY 66 AMQWNTTFHOTLQDPRVGLYFPAGSSSGTVNPVTVPSSISFISRIQDPALNMENI 125
DB 108 AMQWNTTFHOTLQDPRVGLYFPAGSSSGTVNPVTVPSSISFISRIQDPALNMENI 167
QY 126 TSGFGLPLVYLAQGFPLRLITLITPOSIDSWMTSLNFGCTTVCLGONSQSPSTNSHPTS 185
DB 168 TSGFGLPLVYLAQGFPLRLITLITPOSIDSWMTSLNFGCTTVCLGONSQSPSTNSHPTS 227
QY 186 CPPTCGYRWMCRLRFILFLILLCLIFLLVLDYQGLPVCPLIGSSSTSTGPORTC 245
DB 228 CPPTCGYRWMCRLRFILFLILLCLIFLLVLDYQGLPVCPLIGSSSTSTGPORTC 287
QY 246 TTPAGTSMYPSCCCTKPSDGNCTCIPSSWAFGKFLWEMASARFWSLILVFPVQMFV 305
DB 288 MTTAGTSMYPSCCCTKPSDGNCTCIPSSWAFGKFLWEMASARFWSLILVFPVQMFV 347
QY 306 GLSPTVWLSVIMMMYMGPSLYSLISPLPLPIFFCLMWYI 347
DB 348 GLSPTVWLSVIMMMYMGPSLYSLISPLPLPIFFCLMWYI 389

RESULT 8
US-08-360-107A-116
Sequence 116, Application US/08360107A
Patent No. 6017536
GENERAL INFORMATION:
APPLICANT: Bolognesi, Dani P.
APPLICANT: Matthews, Thomas J.
APPLICANT: Wild, Carl T.
APPLICANT: Barney, Shawn O.
APPLICANT: Lambert, Dennis M.
APPLICANT: Petteway, Stephen R.
APPLICANT: Langlois, Alphonse J.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
TITLE OF INVENTION: TRANSMISSION
NUMBER OF SEQUENCES: 149
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/360,107A
FILING DATE: 20-DEC-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-013
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 116:
SEQUENCE CHARACTERISTICS:
LENGTH: 389 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-360-107A-116

Query Match 96.8%; Score 1868; DB 3; Length 389;
Best Local Similarity 98.2%; Pred. No. 5.2e-157;
Matches 336; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 6 GAGAFGLGFTPHHGLGMSPOAGILLETLPANPPASTNRSGROPTPLSPPLRNTHQ 65
DB 48 GAGAFGLGFTPHHGLGMSPOAGILLETLPANPPASTNRSGROPTPLSPPLRNTHQ 107
QY 66 AMQWNTTFHOTLQDPRVGLYFPAGSSSGTVNPVTVPSSISFISRIQDPALNMENI 125
DB 108 AMQWNTTFHOTLQDPRVGLYFPAGSSSGTVNPVTVPSSISFISRIQDPALNMENI 167
QY 126 TSGFGLPLVYLAQGFPLRLITLITPOSIDSWMTSLNFGCTTVCLGONSQSPSTNSHPTS 185
DB 168 TSGFGLPLVYLAQGFPLRLITLITPOSIDSWMTSLNFGCTTVCLGONSQSPSTNSHPTS 227
QY 186 CPPTCGYRWMCRLRFILFLILLCLIFLLVLDYQGLPVCPLIGSSSTSTGPORTC 245
DB 228 CPPTCGYRWMCRLRFILFLILLCLIFLLVLDYQGLPVCPLIGSSSTSTGPORTC 287
QY 246 TTPAGTSMYPSCCCTKPSDGNCTCIPSSWAFGKFLWEMASARFWSLILVFPVQMFV 305
DB 288 MTTAGTSMYPSCCCTKPSDGNCTCIPSSWAFGKFLWEMASARFWSLILVFPVQMFV 347
QY 306 GLSPTVWLSVIMMMYMGPSLYSLISPLPLPIFFCLMWYI 347
DB 348 GLSPTVWLSVIMMMYMGPSLYSLISPLPLPIFFCLMWYI 389

RESULT 9
US-08-484-223B-106
Sequence 106, Application US/08484223B
Patent No. 6020459
GENERAL INFORMATION:
APPLICANT: Bolognesi, Dani P.
APPLICANT: Matthews, Thomas J.
APPLICANT: Wild, Carl T.
APPLICANT: Barney, Shawn O.
APPLICANT: Lambert, Dennis M.
APPLICANT: Petteway, Stephen R.
APPLICANT: Langlois, Alphonse J.
TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
TITLE OF INVENTION: TRANSMISSION
NUMBER OF SEQUENCES: 245
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP

STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,223B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-029
TELEPHONE: (212) 869-9741/8864
TELEFAX: (212) 869-9741/8864
INFORMATION FOR SEQ ID NO: 106:
SEQUENCE CHARACTERISTICS:
LENGTH: 389 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-484-223B-106

Query Match 96.8%; Score 1868; DB 3; Length 389;
Best Local Similarity 98.2%; Pred. No. 5.2e-157;
Matches 336; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 6 GAGAFGLGFTPHGGLGMSPOAGILETLPANPPASTNRSGRQPTPLSPPLANTHQ 65
DB 48 GAGAFGLGFTPHGGLGMSPOAGILETLPANPPASTNRSGRQPTPLSPPLANTHQ 107
QY 66 AMQWNSTTHQTLQDPRVGLYPPAGSSSGTVNPVPTTVSIFSSIFSRIGDPALNMENI 125
DB 108 AMQWNSTTHQTLQDPRVGLYPPAGSSSGTVNPVPTTVSIFSSIFSRIGDPALNMENI 167
QY 126 TSGFLGPLVLVQAGFPLRLITLIPQSLDSWMTSLNPLAGTTVCLGQNSQSPHSPTS 185
DB 168 TSGFLGPLVLVQAGFPLRLITLIPQSLDSWMTSLNPLAGTTVCLGQNSQSPHSPTS 227
QY 186 CPPTCGYRMWCLRRFIIIFLLILCLIFLLVLDYQMLPVCPLIPGSSSTSTGCRFC 245
DB 228 CPPTCGYRMWCLRRFIIIFLLILCLIFLLVLDYQMLPVCPLIPGSSSTSTGCRFC 287
QY 246 TTPAGTSMYSPCCCTKPSDGNCTCIPIPSSMAFGFLMNASARFSLVLVFPQMFV 305
DB 288 MTTAQTSMYSPCCCTKPSDGNCTCIPIPSSMAFGFLMNASARFSLVLVFPQMFV 347
QY 306 GLSPTWVLVIMMMWYGPISLSPFLPLPIFFCLWYI 347
DB 348 GLSPTWVLVIMMMWYGPISLSPFLPLPIFFCLWYI 389

RESULT 10
US-08-919-597-106
Sequence 106, Application US/08919597
Patent No. 6054265
GENERAL INFORMATION:
APPLICANT: Biologues, Dani P.
APPLICANT: Mathews, Thomas J.
APPLICANT: Wild, Carl T.
APPLICANT: Barney, Shawn O.
APPLICANT: Lambert, Dennis M.
APPLICANT: Petteway, Stephen R.
APPLICANT: Langlois, Alphonse J.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION

TITLE OF INVENTION: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
TRANSMISSION
NUMBER OF SEQUENCES: 273
CORRESPONDENCE ADDRESS:
ADDRESSEE: Penie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/919,597
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/470,896
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-020
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
INFORMATION FOR SEQ ID NO: 106:
SEQUENCE CHARACTERISTICS:
LENGTH: 389 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-919-597-106

Query Match 96.8%; Score 1868; DB 3; Length 389;
Best Local Similarity 98.2%; Pred. No. 5.2e-157;
Matches 336; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 6 GAGAFGLGFTPHGGLGMSPOAGILETLPANPPASTNRSGRQPTPLSPPLANTHQ 65
DB 48 GAGAFGLGFTPHGGLGMSPOAGILETLPANPPASTNRSGRQPTPLSPPLANTHQ 107
QY 66 AMQWNSTTHQTLQDPRVGLYPPAGSSSGTVNPVPTTVSIFSSIFSRIGDPALNMENI 125
DB 108 AMQWNSTTHQTLQDPRVGLYPPAGSSSGTVNPVPTTVSIFSSIFSRIGDPALNMENI 167
QY 126 TSGFLGPLVLVQAGFPLRLITLIPQSLDSWMTSLNPLAGTTVCLGQNSQSPHSPTS 185
DB 168 TSGFLGPLVLVQAGFPLRLITLIPQSLDSWMTSLNPLAGTTVCLGQNSQSPHSPTS 227
QY 186 CPPTCGYRMWCLRRFIIIFLLILCLIFLLVLDYQMLPVCPLIPGSSSTSTGCRFC 245
DB 228 CPPTCGYRMWCLRRFIIIFLLILCLIFLLVLDYQMLPVCPLIPGSSSTSTGCRFC 287
QY 246 TTPAGTSMYSPCCCTKPSDGNCTCIPIPSSMAFGFLMNASARFSLVLVFPQMFV 305
DB 288 MTTAQTSMYSPCCCTKPSDGNCTCIPIPSSMAFGFLMNASARFSLVLVFPQMFV 347
QY 306 GLSPTWVLVIMMMWYGPISLSPFLPLPIFFCLWYI 347
DB 348 GLSPTWVLVIMMMWYGPISLSPFLPLPIFFCLWYI 389

RESULT 11
US-08-475-668A-106
Sequence 106, Application US/08475668A
Patent No. 6060065
GENERAL INFORMATION:

APPLICANT: Barney, Shawn O.
APPLICANT: Lambert, Dennis M.
APPLICANT: Petteway, Stephen R.
TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE
TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA VIRUS
TITLE OF INVENTION: TRANSMISSION
NUMBER OF SEQUENCES: 211
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,668A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
INFORMATION FOR SEQ ID NO: 106:
SEQUENCE CHARACTERISTICS:
LENGTH: 389 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-475-668A-106

Query Match 96.8%; Score 1868; DB 3; Length 389;
Best Local Similarity 98.2%; Pred. No. 5.2e-157;
Matches 336; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 6 GAGAFGLFTPHGGLGWSPOAGILFTLPANPPASTNRSGROPTPLSPPLRNTHPQ 65
DB 48 GAGAFGLFTPHGGLGWSPOAGILFTLPANPPASTNRSGROPTPLSPPLRNTHPQ 107
QY 66 AMONNSTTFHOTLQDPVRVGLYFPAGSSSGVTNVPPTVSPISIFSRIGDPAIAMENT 125
DB 108 AMONNSTTFHOTLQDPVRVGLYFPAGSSSGVTNVPPTVSPISIFSRIGDPAIAMENT 167
QY 126 TSGFLGFLVLQAGFLTRILITIPQSLDSWMTSLNFIAGTTVCIGQNSQSPTSNHSPTS 185
DB 168 TSGFLGFLVLQAGFLTRILITIPQSLDSWMTSLNFIAGTTVCIGQNSQSPTSNHSPTS 227
QY 186 CPTPCGYRMNCARRIFLFIILLCLIFLLVLDYQMLPVCPLIPGSSSTTSGPCRTC 245
DB 228 CPTPCGYRMNCARRIFLFIILLCLIFLLVLDYQMLPVCPLIPGSSSTTSGPCRTC 287
QY 246 TTPAOGTSMYSPCCCTKPSDGNCTCIPISPSWAFGKFLMENASARFSLVLVPVQWFEV 305
DB 288 MTTAOGTSMYSPCCCTKPSDGNCTCIPISPSWAFGKFLMENASARFSLVLVPVQWFEV 347
QY 306 GLSPTVWLVSIVMMWYMGPSLSYLSLSPFLPLPIFFCLMVIYI 347
DB 348 GLSPTVWLVSIVMMWYMGPSLSYLSLSPFLPLPIFFCLMVIYI 389

RESULT 12
US-08-485-551A-106
Sequence 106, Application US/08485551A
Patent No. 6068973

GENERAL INFORMATION:
APPLICANT: Bolognesi, Dani P.
APPLICANT: Matthews, Thomas J.
APPLICANT: Wild, Carl T.
APPLICANT: Barney, Shawn O.
APPLICANT: Lambert, Dennis M.
APPLICANT: Petteway, Stephen R.
APPLICANT: Langlois, Alphonse J.
TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE
TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA VIRUS
TITLE OF INVENTION: TRANSMISSION
NUMBER OF SEQUENCES: 211
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,551A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-023
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
INFORMATION FOR SEQ ID NO: 106:
SEQUENCE CHARACTERISTICS:
LENGTH: 389 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-485-551A-106

Query Match 96.8%; Score 1868; DB 3; Length 389;
Best Local Similarity 98.2%; Pred. No. 5.2e-157;
Matches 336; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 6 GAGAFGLFTPHGGLGWSPOAGILFTLPANPPASTNRSGROPTPLSPPLRNTHPQ 65
DB 48 GAGAFGLFTPHGGLGWSPOAGILFTLPANPPASTNRSGROPTPLSPPLRNTHPQ 107
QY 66 AMONNSTTFHOTLQDPVRVGLYFPAGSSSGVTNVPPTVSPISIFSRIGDPAIAMENT 125
DB 108 AMONNSTTFHOTLQDPVRVGLYFPAGSSSGVTNVPPTVSPISIFSRIGDPAIAMENT 167
QY 126 TSGFLGFLVLQAGFLTRILITIPQSLDSWMTSLNFIAGTTVCIGQNSQSPTSNHSPTS 185
DB 168 TSGFLGFLVLQAGFLTRILITIPQSLDSWMTSLNFIAGTTVCIGQNSQSPTSNHSPTS 227
QY 186 CPTPCGYRMNCARRIFLFIILLCLIFLLVLDYQMLPVCPLIPGSSSTTSGPCRTC 245
DB 228 CPTPCGYRMNCARRIFLFIILLCLIFLLVLDYQMLPVCPLIPGSSSTTSGPCRTC 287
QY 246 TTPAOGTSMYSPCCCTKPSDGNCTCIPISPSWAFGKFLMENASARFSLVLVPVQWFEV 305
DB 288 MTTAOGTSMYSPCCCTKPSDGNCTCIPISPSWAFGKFLMENASARFSLVLVPVQWFEV 347
QY 306 GLSPTVWLVSIVMMWYMGPSLSYLSLSPFLPLPIFFCLMVIYI 347
DB 348 GLSPTVWLVSIVMMWYMGPSLSYLSLSPFLPLPIFFCLMVIYI 389

RESULT 13

US-08-471-913A-106
Sequence 106, Application US/08471913A
Patent No. 6093794
GENERAL INFORMATION:
APPLICANT: Bolognesi, Daniel P.
APPLICANT: Matthews, Thomas J.
APPLICANT: Wild, Carl T.
APPLICANT: Barney, Shawn O.
APPLICANT: Lambert, Dennis M.
APPLICANT: Petteway, Stephen R.
APPLICANT: Langlois, Alphonse J.
TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE
TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING EPSTEIN-BARR VIRUS
TITLE OF INVENTION: TRANSMISSION
NUMBER OF SEQUENCES: 214
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edwards LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,913A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-030
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 106:
SEQUENCE CHARACTERISTICS:
LENGTH: 389 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-471-913A-106

Query Match 96.8%; Score 1868; DB 3; Length 389;
Best Local Similarity 98.2%; Pred. No. 5.2e-157;
Matches 336; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 6 GAGAFGLGTPPHGGLGMSPOAGILLETLPANPPASTNRSGROPTPLSPPLRNTHPQ 65
DB 48 GAGAFGLGTPPHGGLGMSPOAGILLETLPANPPASTNRSGROPTPLSPPLRNTHPQ 107

QY 66 AMONNSTTHOTLQDPRVGLYFPAGSSSGTVNVPPTVSPISIFSRIGDPALNMENI 125
DB 108 AMONNSTTHOTLQDPRVGLYFPAGSSSGTVNVPPTVSPISIFSRIGDPALNMENI 167

QY 126 TSGFLGFLVLVQAGFLLTRILITIPQSLDSWMTSLNPLGTTVCLGONSQSPSNHSPTS 185
DB 168 TSGFLGFLVLVQAGFLLTRILITIPQSLDSWMTSLNPLGTTVCLGONSQSPSNHSPTS 227

QY 186 CPPTCGYRWMLCRRTIIFLLCLIFLVLVLDYQGMPLVCPPLIPGSSSTSTGPCRTC 245
DB 228 CPPTCGYRWMLCRRTIIFLLCLIFLVLVLDYQGMPLVCPPLIPGSSSTSTGPCRTC 287

QY 246 TTPAGTSMYPSCCCTKPSDNCCTCIPISSMWAFGKFLMEWASARFSLVLVFPVQMFV 305
DB 288 MTAAGTSMYPSCCCTKPSDNCCTCIPISSMWAFGKFLMEWASARFSLVLVFPVQMFV 347

RESULT 14

US-08-485-264A-106
Sequence 106, Application US/08485264A
Patent No. 6228983
GENERAL INFORMATION:
APPLICANT: Bolognesi, Daniel P.
APPLICANT: Matthews, Thomas J.
APPLICANT: Wild, Carl T.
APPLICANT: Barney, Shawn O.
APPLICANT: Lambert, Dennis M.
APPLICANT: Petteway, Stephen R.
APPLICANT: Langlois, Alphonse J.
TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING
TITLE OF INVENTION: RESPIRATORY SYNCYTIAL VIRUS TRANSMISSION
NUMBER OF SEQUENCES: 232
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edwards LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,264A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-021
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 106:
SEQUENCE CHARACTERISTICS:
LENGTH: 389 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-485-264A-106

Query Match 96.8%; Score 1868; DB 4; Length 389;
Best Local Similarity 98.2%; Pred. No. 5.2e-157;
Matches 336; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 6 GAGAFGLGTPPHGGLGMSPOAGILLETLPANPPASTNRSGROPTPLSPPLRNTHPQ 65
DB 48 GAGAFGLGTPPHGGLGMSPOAGILLETLPANPPASTNRSGROPTPLSPPLRNTHPQ 107

QY 66 AMONNSTTHOTLQDPRVGLYFPAGSSSGTVNVPPTVSPISIFSRIGDPALNMENI 125
DB 108 AMONNSTTHOTLQDPRVGLYFPAGSSSGTVNVPPTVSPISIFSRIGDPALNMENI 167

QY 126 TSGFLGFLVLVQAGFLLTRILITIPQSLDSWMTSLNPLGTTVCLGONSQSPSNHSPTS 185
DB 168 TSGFLGFLVLVQAGFLLTRILITIPQSLDSWMTSLNPLGTTVCLGONSQSPSNHSPTS 227

QY 186 CPPTCGYRWMLCRRTIIFLLCLIFLVLVLDYQGMPLVCPPLIPGSSSTSTGPCRTC 245
DB 228 CPPTCGYRWMLCRRTIIFLLCLIFLVLVLDYQGMPLVCPPLIPGSSSTSTGPCRTC 287

Db 228 CPTCPGVRMCLRRPFIIFLLCLIFLVLVDYQMLPVCPLIPGSSSTTSGPCRTC 287
 QY 246 TTPAGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLMWSARFWSLSLVPVQMFV 305
 Db 288 MTTAGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLMWSARFWSLSLVPVQMFV 347
 QY 306 GLSPTVWLSVIMMMWYMGPSLSYLSLSPFLPLPIFFCLMAYI 347
 Db 348 GLSPTVWLSVIMMMWYMGPSLSYLSLSPFLPLPIFFCLMAYI 389

RESULT 15

US-08-474-349A-106
 ; Sequence 106, Application US/08474349A
 ; Patent No. 6333395

GENERAL INFORMATION:

APPLICANT: Biolognesi, Dani P.
 APPLICANT: Matthews, Thomas J.
 APPLICANT: Wild, Carl T.
 APPLICANT: Barney, Shawn O.
 APPLICANT: Lambert, Dennis M.
 APPLICANT: Pelletway, Stephen R.
 APPLICANT: Langlois, Alphonse J.
 TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE
 TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING HUMAN PARAINFLUENZA
 NUMBER OF INVENTIONS: 517
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Pennie & Edmonds
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/474,349A
 FILING DATE: 07-JUN-1995
 CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Coruzzi, Laura A.
 REGISTRATION NUMBER: 30,742
 REFERENCE/DOCKET NUMBER: 7872-024
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 669-9741/8864
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO. 106:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 389 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 US-08-474-349A-106

Query Match 96.8%; Score 1868; DB 4; Length 389;
 Best Local Similarity 98.2%; Pred. No. 5,2e-157;
 Matches 336; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 6 GAGAVGLGFTPHGSLGMSPOAGILETTPANPPASTNRQSGRQPTPLSPPLRNTHPQ 65
 Db 48 GAGAVGLGFTPHGSLGMSPOAGILETTPANPPASTNRQSGRQPTPLSPPLRNTHPQ 107
 QY 66 AMQNSTTFHQTLDPRVGLYFPAGSSSGTVNPVPTVSPISPSRIGDPAIAMENT 125
 Db 108 AMQNSTTFHQTLDPRVGLYFPAGSSSGTVNPVPTVSPISPSRIGDPAIAMENT 167
 QY 126 TSGFLGLVLVQAGFLLTRILITPQSLDSWMTSLNFIAGTTVCIAGQNSQPTSNHSPTS 165

Db 168 TSGFLGLVLVQAGFLLTRILITPQSLDSWMTSLNFIAGTTVCIAGQNSQPTSNHSPTS 227
 QY 186 CPTCPGVRMCLRRPFIIFLLCLIFLVLVDYQMLPVCPLIPGSSSTTSGPCRTC 245
 Db 228 CPTCPGVRMCLRRPFIIFLLCLIFLVLVDYQMLPVCPLIPGSSSTTSGPCRTC 287
 QY 246 TTPAGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLMWSARFWSLSLVPVQMFV 305
 Db 288 MTTAGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLMWSARFWSLSLVPVQMFV 347
 QY 306 GLSPTVWLSVIMMMWYMGPSLSYLSLSPFLPLPIFFCLMAYI 347
 Db 348 GLSPTVWLSVIMMMWYMGPSLSYLSLSPFLPLPIFFCLMAYI 389

Search completed: February 3, 2003, 09:27:09
 Job time : 23.1585 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd

OM protein - protein search, using sw model

February 3, 2003, 09:15:41 ; Search time 55.6167 Seconds
(without alignments)
831.368 Million cell updates/sec

```

title: US-09-890-752A-1
Perfect score: 1929
Sequence: 1 MGRGDGAGAFGLGFTPPHG.....SLSPFLPLPIFFCLMYI 347

```

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

```
Maximum DB seq length: 0
Maximum DB seq length: 2000000000
```

```

post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

```

A_Gensed 101002:*

1: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1980.DAT.*
2: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1981.DAT.*
3: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1982.DAT.*
4: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1983.DAT.*
5: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1984.DAT.*
6: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1985.DAT.*
7: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1986.DAT.*
8: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1987.DAT.*
9: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1988.DAT.*
10: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1989.DAT.*
11: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1990.DAT.*
12: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1991.DAT.*
13: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1992.DAT.*
14: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1993.DAT.*
15: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1994.DAT.*
16: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1995.DAT.*
17: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1996.DAT.*
18: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1997.DAT.*
19: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1998.DAT.*
20: /SID2/gcgdata/geneeq/geneeqp-emb1/AA1999.DAT.*
21: /SID2/gcgdata/geneeq/geneeqp-emb1/AA2000.DAT.*
22: /SID2/gcgdata/geneeq/geneeqp-emb1/AA2001.DAT.*
23: /SID2/gcgdata/geneeq/geneeqp-emb1/AA2002.DAT.*

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1929	100.0	347	21	AAH10596	HBV fusion protein
2	1879	97.4	389	22	AAG66931	HBV genotype D pre
3	1870	96.9	389	13	AAR27472	1pbas protein. S
4	1870	96.9	389	13	AAR27474	13L promoter/512/c
5	1868	96.8	389	22	AAG62831	Amino acid sequenc
6	1803.5	93.5	393	22	AAU14037	Hepatitis B virus
7	1785	92.4	399	22	AAG66932	HBV genotype B pre
8	1756	91.0	400	17	AAR93801	Hepatitis B virus
9	1752	90.8	389	7	AA60794	Adt-type hepatitis
10	1745	90.5	400	22	AAG66930	HBV genotype C pre

11	1743	90.4	389	22	AAG66916	HBV HBpol protein
12	1735	90.0	393	9	AAP80419	Sequence of adr ty
13	1727	89.5	389	15	AAR35287	Deduced sequence c
14	1721	88.2	400	21	AAI54045	Amino acid sequencu
15	1713	88.8	388	12	AAAT0851	Modified hepatitis
16	1704	88.3	389	23	AAAM52682	Hepatitis B virus
17	1704	88.3	400	21	AAAY44349	Human hepatitis B
18	1700	88.1	400	22	AAG66928	HBV genotype A pr
19	1686.5	87.4	387	12	AAAT0852	Modified hepatitic
20	1684	87.3	400	22	AAG66939	HBV genotype B pr
21	1587	82.3	400	22	AAG66933	HBV genotype F pre
22	1539	79.8	281	13	AAR377411	spsag protease. Sy
23	1534	79.5	281	20	AAAY32835	HBsAg Presz-S regl
24	1498.5	77.7	286	19	AAAM44780	Modified hepatitic
25	1469	76.2	281	7	AAP60163	Subtype adr hepati
26	1469	76.2	281	7	AAP60560	Hepatitis B virus
27	1469	76.2	281	7	AAP60561	C-terminal detecti
28	1465	75.9	281	8	AAP70294	Subtype adr HBsAg
29	1460	75.7	281	15	AAR62870	Hepatitis B virus
30	1460	75.7	281	22	AAAB48419	HBsAg Presz-S doma
31	1444	74.9	281	22	AAAB2620	Protein #3 used to
32	1437	74.5	281	20	AAAY32834	HBsAg Presz-S regl
33	1436	74.4	281	23	AAAS52693	HBsAg Presz-S regl
34	1428	74.0	281	15	AARS5281	Hepatitis B virus
35	1418	73.5	281	7	AAP60561	Deduced sequence o
36	1384	71.7	281	7	AAP50164	Hepatitis B virus
37	1279	66.3	249	21	AAI52585	Subtype adw hepati
38	1265	65.6	246	22	AAAG62930	Pan DR epitope/HBV
39	1264	65.5	269	22	AAAH8418	Amino acid sequenc
40	1262.5	65.4	280	12	AAAT10850	Protein #2 used to
41	1260	65.3	246	19	AAAG62927	Modified hepatitis
42	1259	65.3	246	23	AAE19897	Hepatitis B virus
43	1252	64.9	246	2	AAPI0001	Hepatitis B virus
44	1252	64.9	226	18	AAAD0945	Sequence encoded b
45	1249	64.7	226	2	ABE77764	Hepatitis B virus
						Amino acid sequenc

ALIGNMENTS

XX	AA10596	
ID	AA10596 standard; Protein; 347 AA.	
AC	AA10596;	
DT	08-JAN-2001 (first entry)	
DE	HEV fusion protein comprising LHB and RGD.	
KW	Fusion protein; protein coat; virus-specific packaging signal; psi	
KX	virus protein; cell permeability; cell-specific binding site; LHB;	
KX	large surface protein; core antigen; gene therapy.	
OS	Hepatitis b virus.	
OS	Synthetic.	
PN	MO200046376-A2.	
PD	10-AUG-2000.	
PX		
PF	04-FEB-2000; 2000WC-DE00163.	
PX		
PR	05-FEB-1999; 99DE-1004800.	
PA	(HILD/) HILDT E.	
PX		
PI	Hildt E, Hofschneider P;	
PX		
DR	MPI; 2000-514959/46.	
DR	N-PSDB; AAA71734.	
PT	Particle for cell-specific gene delivery, useful in gene therapy.	

comprises nucleic acid in protein coat that includes a fusion protein of viral protein, permeability peptide and cell-binding site -

Claim 14; Fig 1; 34pp; German.

This invention describes a novel particle (A), comprising a protein coat with a fusion protein (FP), and, inside the coat, a nucleic acid (I) including the sequence for a virus-specific packaging signal (psi) and a structural gene. FP contains a virus protein (VP), a peptide (P) that mediates cell permeability and a heterologous cell-specific binding site (RBP). The invention also describes (1) producing (A) in which FP contains an LHBs (large surface protein of hepatitis B virus (HBV)) and (2) preparing (A) in which FP contains an HBV core antigen (HBcAg), (P) and RBP; (3) FP; (4) DNA encoding FP; and (5) expression vector containing the DNA of (d). The products of the invention are used in gene therapy of cells and tissues, in vivo or ex vivo. This sequence represents a fusion protein which is described in the method of the invention.

Sequence 347 AA;
Query Match 100.0%; Score 1929; DB 21; Length 347;
Best Local Similarity 100.0%; Pred. No. 9.2e-143; Indels 0; Gaps 0;
Matches 347; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MGRGDAAGAGLGFPPHGLGWSPOAGIILETLPANPPASTNRSGRQPTPLSPUR 60
1 MGRGDAAGAGLGFPPHGLGWSPOAGIILETLPANPPASTNRSGRQPTPLSPUR 60
61 NTHPQAMQWNTTFHQTLDPRVRGLYPFAGSSSGTNVPVTTSPISISRIQDPL 120
61 NTHPQAMQWNTTFHQTLDPRVRGLYPFAGSSSGTNVPVTTSPISISRIQDPL 120
121 NMENITSGFLGPLVLOAGFFLTRILITPQSIDSWTSLNFGGTTVCLGQNSGPTSN 180
121 NMENITSGFLGPLVLOAGFFLTRILITPQSIDSWTSLNFGGTTVCLGQNSGPTSN 180
181 HSPFSCPTCPGYRMWMLRRFIFLFLILCLIFLLVLDYQMLPVCPLIRGSSSTTSG 240
181 HSPFSCPTCPGYRMWMLRRFIFLFLILCLIFLLVLDYQMLPVCPLIRGSSSTTSG 240
241 PCRTCTTTPAGTSMYRSCCTKPSDNGCTCIPSSWAFGKFLMWSARFSLSLVFP 300
241 PCRTCTTTPAGTSMYRSCCTKPSDNGCTCIPSSWAFGKFLMWSARFSLSLVFP 300
301 VQWFGSLPTWLSVIMMMWYMGPSLSLSPLPLPIFFCLMVTYI 347
301 VQWFGSLPTWLSVIMMMWYMGPSLSLSPLPLPIFFCLMVTYI 347

RESULT 2
AAG66931
ID AAG66931 standard; Protein: 389 AA.

AC AAG66931;
19-OCT-2001 (first entry)
HBV genotype D pres1/pres2/HBcAg polypeptide.
HBV genotype D pres1/pres2/HBcAg polypeptide.
Hepatitis B virus; HBV, preCore; Core; pres1; pres2; HBs; HBx; HBp1;
HBcAg; anti-viral; vaccine; genotype G; genotyping; HBcAg; HBcAg;
Hepatitis B virus.
W0200140279-A2.
07-JUN-2001.
20-NOV-2000, 2000WO-EP11526.
93-DEC-1999; 99BP-0870252.
07-DEC-1999; 99US-0169287.

(INNO-) INNOGENETICS NV.
Stuyver L, Van Geyt C, De Gendt S;
WPI; 2001-374785/39.

Novel isolated and/or purified hepatitis B virus polypeptide and polynucleotide sequences that are phylogenetically different from HBV genotype A-F molecules, useful for HBV diagnosis, prophylaxis and therapy -

Example 3; Fig 6; 94pp; English.

The invention relates to the complete nucleic acid sequence of a new human hepatitis B virus (HBV) genotype, provisionally named genotype G. This genotype was found with a high prevalence in patients from Europe and the USA. The chronically infected with HBV and residing in Europe and the USA. The invention relates to a fully defined sequence of 3248 nucleotides as given in specification, a sequence with 92% identity to the given sequence, or sequence that is degenerate to the mentioned sequences. These polynucleotides are useful for detecting antibodies in encoded by the polynucleotides are useful for detecting antibodies in a biological sample. Ligands that bind to the proteins and antibodies directed against the proteins are useful for detecting the proteins. They are also useful for preparing a vaccine or medicament for treating HBV infections. The present sequence is provided in an amino acid sequence alignment of the pres1, pres2 and HBcAg open reading frame of the different HBV genotypes.

Sequence 389 AA;
Query Match 97.4%; Score 1879; DB 22; Length 389;
Best Local Similarity 98.5%; Pred. No. 8.4e-139;
Matches 337; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

6 GAGAFGLGFTPPHGLGWSPOAGIILETLPANPPASTNRSGRQPTPLSPURTHPQ 65
6 GAGAFGLGFTPPHGLGWSPOAGIILETLPANPPASTNRSGRQPTPLSPURTHPQ 107
48 GAGAFGLGFTPPHGLGWSPOAGIILETLPANPPASTNRSGRQPTPLSPURTHPQ 107
66 AMQWNTTFHQTLDPRVRGLYPFAGSSSGTNVPVTTSPISISRIQDPLNMENI 125
108 AMQWNTTFHQTLDPRVRGLYPFAGSSSGTNVPVTTSPISISRIQDPLNMENI 167
126 TSGFLGPLVLOAGFFLTRILITPQSIDSWTSLNFGGTTVCLGQNSGPTSNHSP 185
168 TSGFLGPLVLOAGFFLTRILITPQSIDSWTSLNFGGTTVCLGQNSGPTSNHSP 227
186 CPPTCPGYRMWMLRRFIFLFLILCLIFLLVLDYQMLPVCPLIRGSSSTTSGPRTC 245
228 CPPTCPGYRMWMLRRFIFLFLILCLIFLLVLDYQMLPVCPLIRGSSSTTSGPRTC 287
246 TTPAGTSMYRSCCTKPSDNGCTCIPSSWAFGKFLMWSARFSLSLVFPQWVF 305
288 TTPAGTSMYRSCCTKPSDNGCTCIPSSWAFGKFLMWSARFSLSLVFPQWVF 347
306 GLSPTWLSVIMMMWYMGPSLSLSPLPLPIFFCLMVTYI 347
348 GLSPTWLSVIMMMWYMGPSLSLSPLPLPIFFCLMVTYI 389

RESULT 3
AAR27472
ID AAR27472 standard; Protein: 389 AA.

AC AAR27472;
24-FEB-1993 (first entry)
HBcAg protein.
Hepatitis B virus; HBV, M protein; small pre-S antigen; sHBsAg;

KM vaccinia virus; hemorrhagic region; u. promoter; NYVAC; recombinant;
 KM HBV L; large pre-S antigen; 1psag; fusion protein; pre-S region;
 KM S12/core; S1; S2; Copenhagen vaccine strain; vaccinia virus;
 KM virulence factor; deletion loci; recipient loci.
 OS Synthetic.
 FH Key
 FH region 1.108
 FT /label S1
 FT 109.163
 FT /label= S2
 FT misc_RNA 164.389
 FT /label= S
 PN MO9215672-A.
 XX 17-SEP-1992.
 PD
 XX 09-MAR-1992; 92MO-US01906.
 PF
 XX 07-MAR-1991; 91US-0666056.
 PR 11-JUN-1991; 91US-0713967.
 PR 06-MAR-1992; 92US-0847951.
 XX
 PA (VIRO-) VIROGENETICS CORP.
 XX
 PI Cox WI, De Taisne C, Francis J, Gettig RR, Johnson GP,
 PI Limbach KJ, Norton RK, Paolietti E, Perkus ME, Pincus SE;
 PI Riviere M, Tartaglia J, Taylor J;
 DR WPI; 1992-331718/40.
 XX
 DR N-PSDB; AAQ29104.
 XX
 PT Vaccine comprises recombinant, attenuated pox-virus - use for
 PT vaccinating against viral infections such as rabies, hepatitis B,
 PT HIV, HSV, EBV, CMV, mumps etc.
 PS Disclosure; Fig 11; 456pp; English.
 XX
 PS The sequence given is encoded by an expression vector which comprises
 CC the hepatitis B virus (HBV) L protein (large pre-S antigen, 1psag)
 CC gene linked to the comox hemorrhagic region (u) promoter. This DNA
 CC sequence was used in the construction of a NYVAC recombinant
 CC expressing the HBV gene. Other HBV genes were also used in the
 CC construction. These were HBV M protein (small pre-S antigen, spsag)
 CC and a fusion protein composed of the entire pre-S region (S12/core,
 CC S1 + S2). Each of these gene sequences were inserted individually
 CC into three different sites of NYVAC separated by from each other by
 CC large regions of vaccinia DNA containing essential genes. NYVAC is a
 CC Copenhagen vaccine strain of vaccinia virus which has been modified by
 CC deletion of six non-essential regions of the genome encoding known or
 CC potential virulence factors. The deletion loci were engineered as
 CC recipient loci for the insertion of foreign genes. The spacing of the
 CC three inserted sequences ensured that any recombination that did occur
 CC would lead to disruption of the vaccinia genome and would cause
 CC unviable vaccinia virus. See also AAQ35501-864.
 CC
 XX
 SQ Sequence 389 AA;
 Query Match 96.94; Score 1870; DB 13; Length 389;
 Best Local Similarity 98.04; Pred. No. 4.2e-138;
 Matches 335; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 QY 6 GAGAFGLGFTPHGGLGMSPOAGIILETLPANPPASTNRQSGROTPPLSPRLNTHPQ 65
 DB 48 GAGAFGLGFTPHGGLGMSPOAGIILETLPANPPASTNRQSGROTPPLSPRLNTHPQ 107
 QY 66 AMONNSTTFHQTQDPRVRGLYPAGSSSGTVVPPTVSPISISFRIQDPLNMENT 125
 DB 108 AMONNSTTFHQTQDPRVRGLYPAGSSSGAVVPPTTASPLSISFRIQDPLNMENT 167
 QY 126 TSGFLGFLVLQAGFLTLTLITPQSLDSWMTSLNPLGCTVCLGQNSQSPTSNHSPTS 185

DB 168 TSGFLGFLVLQAGFLTLTLITPQSLDSWMTSLNPLGCTVCLGQNSQSPTSNHSPTS 227
 QY 186 CPPTCPGYRMWMLRRFIFLFLILLCLIFLVLVDYQGLPVCPLIPSGSTSTGRCRT 245
 DB 228 CPPTCPGYRMWMLRRFIFLFLILLCLIFLVLVDYQGLPVCPLIPSGSTSTGRCRT 287
 QY 246 TTPAGTSMYPSCCCTKPSDNGCTCIPPSWAEGKFLMENASARFSLILVFPVQMFV 305
 DB 288 MTTAQSTSMYPSCCCTKPSDNGCTCIPPSWAEGKFLMENASARFSLILVFPVQMFV 347
 QY 306 GUSPTVWLSVIMMMYMGPSLSYSLSPFLPLIPFCIMVYI 347
 DB 348 GUSPTVWLSVIMMMYMGPSLSYSLSPFLPLIPFCIMVYI 389
 RESULT 4
 AAR27474
 ID AAR27474 standard; Protein; 389 AA.
 XX
 AC AAR27474;
 XX
 DT 24-FEB-1993 (first entry)
 XX
 DE 13L promoter/S12/core gene.
 XX
 KW Hepatitis B virus; HBV, M protein; small pre-S antigen; spsag;
 KW vaccinia virus; Amara moorei entomopoxvirus; AmEPV; 42 kD; promoter;
 KM NYVAC; recombinant; HBV L; large pre-S antigen; 1psag; fusion protein;
 KM pre-S region; S12/core; S1; S2; Copenhagen vaccine strain;
 KM vaccinia virus; virulence factor; deletion loci; recipient loci.
 XX
 OS Synthetic.
 XX
 FH Key
 FH region 1.108
 FT /label= S1
 FT 109.163
 FT /label= S2
 FT 164.389
 FT /label= S
 PT Region
 PT
 PN MO9215672-A.
 XX
 PD 17-SEP-1992.
 XX
 PF 09-MAR-1992; 92MO-US01906.
 PR 07-MAR-1991; 91US-0666056.
 PR 11-JUN-1991; 91US-0713967.
 PR 06-MAR-1992; 92US-0847951.
 XX
 PA (VIRO-) VIROGENETICS CORP.
 XX
 PI Cox WI, De Taisne C, Francis J, Gettig RR, Johnson GP,
 PI Limbach KJ, Norton RK, Paolietti E, Perkus ME, Pincus SE;
 PI Riviere M, Tartaglia J, Taylor J;
 DR WPI; 1992-331718/40.
 DR N-PSDB; AAQ29106.
 XX
 PT Vaccine comprises recombinant, attenuated pox-virus - use for
 PT vaccinating against viral infections such as rabies, hepatitis B,
 PT HIV, HSV, EBV, CMV, mumps etc.
 PS Disclosure; Fig 15; 456pp; English.
 XX
 PS The sequence given is encoded by an expression cassette which
 CC comprises the hepatitis B virus (HBV) L protein (large pre-S antigen,
 CC 1psag) gene which is precisely linked to the Amara moorei
 CC entomopoxvirus (AmEPV) 42 kD promoter. This DNA sequence was used in
 CC the construction of a NYVAC recombinant expressing the HBV gene.
 CC Other HBV genes were also used in the construction. These were HBV M

XX 09-JUL-1999; 99US-0350841.
 PR (TRIM-) TRIMERIS INC.
 PA
 XX
 PI Jeffs P, Lackey JW, Erickson JB, Lawless MK, Merutka G;
 DR WPI; 2001-442157/447.
 XX
 PT Identifying a compound that inhibits the formation of or disrupts a
 PT DP107/DP178 complex, especially compounds with antitumor, antiviral
 PT or intracellular modulatory activity, by detecting the formation of a
 PT DP107/DP178 complex -
 XX
 PS Disclosure; Fig 35; 259pp; English.
 CC The present invention relates to peptides which exhibit anti-retroviral
 CC activity. The peptides of the invention (AAU2559-AAU14009) comprise
 CC DP178-like and DP107-like peptides. The DP178 peptide corresponds
 CC to amino acids 639-673 of the transmembrane protein gp41 from human
 CC immunodeficiency virus 1 (HIV-1) isolate LAI. The DP107 peptide
 CC corresponds to amino acids 558-595 of gp41 from HIV-1LAI. The invention
 CC also relates to a method of identifying compounds that inhibit the
 CC formation of or disrupts a DP107/DP178 complex. The method comprises
 CC detecting the formation of a DP107/DP178 complex, both in the presence
 CC or absence of a test compound, in a reaction mixture containing DP107
 CC and DP178 peptides. The method is useful for identifying compounds,
 CC including small molecule compounds, which may themselves exhibit
 CC antitumor, antiviral or intracellular modulatory activity. The
 CC DP178-like/DP107-like peptides are useful to inhibit human and non-human
 CC retroviral, particularly HIV, transmission to uninfected cells. The
 CC present sequence represents a peptide sequence from Hepatitis B virus
 CC subtype AYw major surface antigen precursor S.
 CC
 SQ Sequence 393 AA;
 Query Match 93.5%; Score 1803.5; DB 22; Length 393;
 Best Local Similarity 96.5%; Pred. No. 6.7e-133;
 Matches 333; Conservative 2; Mismatches 7; Indels 3; Gaps 3;
 QY 6 GAGAFGLGTPPP-HGGLGMSPOAGILETLTPANPPASTNROSGRQPTPLSPPLRNT-H 63
 DB 49 GAGAFGLGTPPPHGGGLGMSPOAGILETLTPANPPASTNROSGRQPTPLSPPLRNTII 108
 QY 64 PAMQNSSTTF-HQTLQDPRVGLVPPAGSSSGTVNPPTTVSIFSIISFGIPALNM 122
 DB 109 PAMQNSSTTFHQLQDPRVGLVPPAGSSSGTVNPPTTVSIFSIISFGIPALNM 168
 QY 123 ENITSGFLPPLVLAQGFLLTRILITIPQSLDSWMTSLNFGTTCVCGONSQSTSHS 182
 DB 169 ENITSGFLPPLVLAQGFLLTRILITIPQSLDSWMTSLNFGTTCVCGONSQSTSHS 228
 QY 183 PTCSCPTCGVYMWCLRRPIIFLLILCLIFLLVLDYQGLPVCPLIPGSSSTTGPC 242
 DB 229 PTCSCPTCGVYMWCLRRPIIFLLILCLIFLLVLDYQGLPVCPLIPGSSSTTGPC 288
 QY 243 RTCTTAAQTSWYPSCCCTKSDGNCCTCIPSSWAFGKFLMNASAFSWLSILVPPVQ 302
 DB 289 RTCTTAAQTSWYPSCCCTKSDGNCCTCIPSSWAFGKFLMNASAFSWLSILVPPVQ 348
 QY 303 WTVGLSPVWMLSVIWMWYWGBSLYSILSPFLPLPIFPCLMVYI 347
 DB 349 WTVGLSPVWMLSVIWMWYWGBSLYSILSPFLPLPIFPCLMVYI 393

DE HBV genotype E pres1/pres2/HBsAg polypeptide.
 XX Hepatitis B virus; HBV; preCore; Core; pres1; pres2; HBs; HBx; HBp1,
 KW HBsAg; antiviral; vaccine; genotype G; genotyping; HBeAg; HBeAg.
 OS Hepatitis B virus.
 XX
 PN WO200140279-A2.
 XX
 PD 07-JUN-2001.
 XX
 PF 20-NOV-2000; 2000WO-BP11526.
 XX
 PR 03-DEC-1999; 99EP-0870252.
 PR 07-DEC-1999; 99US-0169287.
 XX
 PA (INNO-) INNOGENETICS NV.
 PI Stuyver L, Van Geyt C, De Gendt S;
 DR WPI; 2001-374785/39.
 XX
 PT Novel isolated and/or purified hepatitis B virus polypeptide and
 PT polynucleotide sequences that are phylogenetically different from HBV
 PT genotype A-F molecules, useful for HBV diagnosis, prophylaxis and
 PT therapy -
 XX
 PS Example 3; Fig 6; 94pp; English.
 CC The invention relates to the complete nucleic acid sequence of a new
 CC human hepatitis B virus (HBV) genotype, provisionally named genotype
 CC G. This genotype was found with a high prevalence in patients
 CC chronically infected with HBV and residing in Europe and the USA. The
 CC invention relates to a fully defined sequence of 3248 nucleotides as
 CC given in specification, a sequence that is degenerate to the given
 CC sequence, or sequence that is degenerate to the mentioned sequences.
 CC These polynucleotides are useful for HBV genotyping. The proteins
 CC encoded by the polynucleotides are useful for detecting antibodies in
 CC a biological sample. Ligands that bind to the proteins and antibodies
 CC directed against the proteins are useful for detecting the proteins
 CC and for detecting HBsAg and HBeAg (precore precursor proteins). They
 CC are also useful for preparing a vaccine or medicament for treating
 CC HBV infections. The present sequence is provided in an amino acid
 CC sequence alignment of the pres1, pres2 and HBsAg open reading frame
 CC of the different HBV genotypes.
 CC
 SQ Sequence 399 AA;
 Query Match 92.4%; Score 1783; DB 22; Length 399;
 Best Local Similarity 93.0%; Pred. No. 2.7e-131;
 Matches 318; Conservative 11; Mismatches 13; Indels 0; Gaps 0;
 QY 6 GAGAFGLGTPPPHGGGLGMSPOAGILETLTPANPPASTNROSGRQPTPLSPPLRNT-HQ 65
 DB 58 GAGAFGLGTPPPHGGGLGMSPOAGILETLTPANPPASTNROSGRQPTPLSPPLRNT-HQ 117
 QY 66 AMQNSSTTFHQLQDPRVGLVPPAGSSSGTVNPPTTVSIFSIISFGIPALNMENI 125
 DB 118 AMQNSSTTFHQLQDPRVGLVPPAGSSSGTVNPPTTVSIFSIISFGIPALNMENI 177
 QY 126 TSGLFLPPLVLAQGFLLTRILITIPQSLDSWMTSLNFGTTCVCGONSQSTSHSPTS 185
 DB 178 TSGLFLPPLVLAQGFLLTRILITIPQSLDSWMTSLNFGTTCVCGONSQSTSHSPTS 237
 QY 186 CPPTCGVYMWCLRRPIIFLLILCLIFLLVLDYQGLPVCPLIPGSSSTTGPCRTC 245
 DB 238 CPPTCGVYMWCLRRPIIFLLILCLIFLLVLDYQGLPVCPLIPGSSSTTGPCRTC 297
 QY 246 TTPAQTSMYPSCCCTKSDGNCCTCIPSSWAFGKFLMNASAFSWLSILVPPVQFV 305
 DB 298 TTPAQTSMYPSCCCTKSDGNCCTCIPSSWAFGKFLMNASAFSWLSILVPPVQFV 357
 QY 306 GLSPVWMLSVIWMWYWGBSLYSILSPFLPLPIFPCLMVYI 347

Db 358 GLSPTVWLSVIMMMWYMGPSLXIILSPFLPIFCLMWYI 399

RESULT 8

AA93801
ID AAF93801 standard; Protein; 400 AA.

AC AAF93801;

DT 22-OCT-1996 (first entry)

DE Hepatitis B virus pre S1/pre S2/S protein.

HE Hepatitis C virus; HCV; hepatitis B virus; HBV; fusion protein;

KM Immunisation; vaccine; infection; core protein; adr-1.

OS Hepatitis B virus.

PN MO9610997-A1.

PD 18-APR-1996.

PF 05-OCT-1995; 95MO-US13552.

PR 06-JUN-1995; 95US-0467859.

PR 05-OCT-1994; 94US-0318248.

PA (APOL-) APOLLON INC.

PA (GEHO) GEN HOSPITAL CORP.

PI Coney LR, Pachuk CJ, Tokushige K, Wakita T, Wands J;

PI Zurawski VR;

DR WPI; 1996-209642/21.

DR N-PSDB; AAT36611.

XX Nucleic acid encoding hepatitis B and C virus fusion proteins - or

XX incomplete hepatitis C virus genome, are useful in vaccines for

XX prevention or treatment of HBV and HCV infections

XX Example 1; Page 32-34; 53pp; English.

XX A new fusion protein comprises a hepatitis B virus (HBV) S gene

XX protein (AAT36611) coupled to amino acids 1-69, 1-70 or 1-134 of the

XX hepatitis C virus (HCV) core protein (AAT28348).

XX The nucleic acid has the coding sequence linked to a CMV promoter,

XX CC RSV enhancer, polyadenylation sequence and opt. the 5'-UTR of HCV.

XX CC

XX Sequence 400 AA;

XX

XX

XX

XX

XX

XX

XX

XX

Db 299 TTPAGTSMFSCCTKRPDGNCTCIPSSMAFARFLMEWAVRFSWLSILVPPVQWFA 358

Qy 306 GLSPTVWLSVIMMMWYMGPSLXIILSPFLPIFCLMWYI 347

Db 359 GLSPTVWLSVIMMMWYMGPSLXIILSPFLPIFCLMWYI 400

RESULT 9

AA60794
ID AAF60794 standard; Protein; 389 AA.

AC AAF60794;

DT 30-OCT-1991 (first entry)

DE Adr-type hepatitis B virus SA, intermediate SA and large SA.

KM Surface antigen; vaccine; hugK-14.

OS Adr-type Hepatitis B virus.

PN MO8603975-A.

PD 17-JUL-1986.

PF 28-DEC-1985; 85MO-JP00733.

PR 28-DEC-1984; 84UP-0274592.

PA (NICA-) JAPAN FOUND CANCER.

PA (MEIP) MEIJI MILK PROD CO LTD.

PA (GANK-) GAN KENKYU KAI ZH.

PI KOIKE K, SUGANO H, ODA M, KATAYANAGI S;

PI WPI; 1986-196794/30.

DR N-PSDB; AAN60714.

XX The sequence allows for the preparation of adr-type hepatitis B surface

XX antigens for vaccine, without the use of human blood. The proteins may be

XX expressed from human hepatoma cell strain hugK-14.

XX Disclousure; Fig 5; 30pp; Japanese.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Db 299 TTPAGTSMFSCCTKRPDGNCTCIPSSMAFARFLMEWAVRFSWLSILVPPVQWFA 358

Qy 306 GLSPTVWLSVIMMMWYMGPSLXIILSPFLPIFCLMWYI 347

Db 359 GLSPTVWLSVIMMMWYMGPSLXIILSPFLPIFCLMWYI 400

RESULT 9

AA60794
ID AAF60794 standard; Protein; 389 AA.

AC AAF60794;

DT 30-OCT-1991 (first entry)

DE Adr-type hepatitis B virus SA, intermediate SA and large SA.

KM Surface antigen; vaccine; hugK-14.

OS Adr-type Hepatitis B virus.

PN MO8603975-A.

PD 17-JUL-1986.

PF 28-DEC-1985; 85MO-JP00733.

PR 28-DEC-1984; 84UP-0274592.

PA (NICA-) JAPAN FOUND CANCER.

PA (MEIP) MEIJI MILK PROD CO LTD.

PA (GANK-) GAN KENKYU KAI ZH.

PI KOIKE K, SUGANO H, ODA M, KATAYANAGI S;

PI WPI; 1986-196794/30.

DR N-PSDB; AAN60714.

XX The sequence allows for the preparation of adr-type hepatitis B surface

XX antigens for vaccine, without the use of human blood. The proteins may be

XX expressed from human hepatoma cell strain hugK-14.

XX Disclousure; Fig 5; 30pp; Japanese.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

QY 186 CPTTCGVRMNCIRRFIIIFLLICLIFLLVLDYQMLFVCPILIPSSSTTSGPCRTC 245
 DB 228 CPICPGVRMNCIRRFIIIFLLICLIFLLVLDYQMLFVCPILIPGTSITSGPCRTC 287
 QY 246 TTPAQGTSMYPSCCCTKPSDNCCTCIPSSMARCKFLMEASARFWSLLVFPVQMFV 305
 DB 288 TIPAGTSMYPSCCCTKPSDNCCTCIPSSMARCKFLMEASARFWSLLVFPVQMFV 347
 QY 306 GLSPTVWLTVIMMMWYMGPSLYSLSPPLPLPIFFCLMVIYI 347
 DB 348 GLSPTVWLTVIMMMWYMGPSLYSLSPPLPLPIFFCLMVIYI 389

RESULT 10

AAG66930
 ID AAG66930 standard; Protein; 400 AA.

AC AAG66930;
 DT 19-OCT-2001 (first entry)
 XX

DE HBV genotype C pres1/pres2/HBsAg polypeptide.

KM Hepatitis B virus; HBV; preCore; Core; pres1; pres2; HBS; HBx; HBp1;
 KM HBsAg; antiviral; vaccine; genotype G; genotyping; HBcAg; HBeAg.

OS Hepatitis B virus.

PN WO200140279-A2.

PD 07-JUN-2001.

PF 20-NOV-2000; 2000WO-EP11526.

PR 03-DEC-1999; 99EP-0870252.

PR 07-DEC-1999; 99US-0169287.

PA (INNO-) INNOGENETICS NV.

PI Stuyver L, Van Geyt C, De Gendt S;

DR WPI; 2001-374785/39.

PT Novel isolated and/or purified hepatitis B virus polypeptide and
 PT polynucleotide sequences that are phylogenetically different from HBV
 PT genotype A-F molecules, useful for HBV diagnosis, prophylaxis and
 therapy

XX Example 3; Fig 6; 94p; English.

CC The invention relates to the complete nucleic acid sequence of a new
 CC human hepatitis B virus (HBV) genotype, provisionally named genotype
 CC G. This genotype was found with a high prevalence in patients
 CC chronically infected with HBV and residing in Europe and the USA. The
 CC invention relates to a fully defined sequence of 3248 nucleotides as
 CC given in specification, a sequence with 92% identity to the given
 CC sequence, or sequence that is degenerate to the mentioned sequences.
 CC These polynucleotides are useful for HBV genotyping. The proteins
 CC encoded by the polynucleotides are useful for detecting antibodies in
 CC a biological sample. Ligands that bind to the proteins and antibodies
 CC directed against the proteins are useful for detecting the proteins
 CC and for detecting HBcAg and HBeAg (precore precursor proteins). They
 CC are also useful for preparing a vaccine or medication for treating
 CC HBV infections. The present sequence is provided in an amino acid
 CC sequence alignment of the pres1, pres2 and HBsAg open reading frame
 CC of the different HBV genotypes.

CC Sequence 400 AA;

Query Match 90.5%; Score 1745; DB 22; Length 400;
 Best Local Similarity 90.9%; Pred. No. 2.6e-128;
 Matches 311; Conservative 11; Mismatches 20; Indels 0; Gaps 0;

QY 6 GAGAFGLGTTPPHGILGNSPOAGILETLTPANPPASINRSGROPTPLSPPLNTHPQ 65
 DB 59 GAGAFGPPTPHGILGNSPOAGVLTVPVAPPASTNRSGRPDIPSPPLADSHQ 118
 QY 66 AMQNSTTFHQLADPRVGLFEPAGSSGSSGVNVPPTVASISIFRIGDPALMENI 125
 DB 119 AMQNSTTHQLADPRVGLFEPAGSSGSSGVNVPPTVASISISRTGDPALMENI 178
 QY 126 TSGFLGPLLVLAGFELLTRILITIPQSLDSWTSINFLGAGTTCVCGONSOSPSTSHSPTS 185
 DB 179 TSGFLGPLLVLAGFELLTRILITIPQSLDSWTSINFLGAGTTCVCGONSOSPSTSHSPTS 238
 QY 186 CPTTCGVRMNCIRRFIIIFLLICLIFLLVLDYQMLFVCPILIPSSSTTSGPCRTC 245
 DB 239 CPICPGVRMNCIRRFIIIFLLICLIFLLVLDYQMLFVCPILIPGTSITSGPCRTC 298
 QY 246 TTPAQGTSMYPSCCCTKPSDNCCTCIPSSMARCKFLMEASARFWSLLVFPVQMFV 305
 DB 299 TIPAGTSMYPSCCCTKPSDNCCTCIPSSMARCKFLMEASARFWSLLVFPVQMFV 358
 QY 306 GLSPTVWLTVIMMMWYMGPSLYSLSPPLPLPIFFCLMVIYI 347
 DB 359 GLSPTVWLTVIMMMWYMGPSLYSLSPPLPLPIFFCLMVIYI 400

RESULT 11

AAG66919
 ID AAG66919 standard; Protein; 399 AA.

AC AAG66919;

DT 19-OCT-2001 (first entry)

DE HBV HBp1 protein.

KM Hepatitis B virus; HBV; preCore; Core; pres1; pres2; HBS; HBx; HBp1;
 KM HBsAg; antiviral; vaccine; genotype G; genotyping; HBcAg; HBeAg.

OS Hepatitis B virus.

PN WO200140279-A2.

PD 07-JUN-2001.

PF 20-NOV-2000; 2000WO-EP11526.

PR 03-DEC-1999; 99EP-0870252.

PR 07-DEC-1999; 99US-0169287.

PA (INNO-) INNOGENETICS NV.

PI Stuyver L, Van Geyt C, De Gendt S;

DR WPI; 2001-374785/39.

DR N-PSDB; AAH77562.

PT Novel isolated and/or purified hepatitis B virus polypeptide and
 PT polynucleotide sequences that are phylogenetically different from HBV
 PT genotype A-F molecules, useful for HBV diagnosis, prophylaxis and
 therapy

XX Claim 16; Fig 1; 94p; English.

CC The invention relates to the complete nucleic acid sequence of a new
 CC human hepatitis B virus (HBV) genotype, provisionally named genotype
 CC G. This genotype was found with a high prevalence in patients
 CC chronically infected with HBV and residing in Europe and the USA. The
 CC invention relates to a fully defined sequence of 3248 nucleotides as
 CC given in specification, a sequence with 92% identity to the given
 CC sequence, or sequence that is degenerate to the mentioned sequences.
 CC These polynucleotides are useful for HBV genotyping. The proteins
 CC encoded by the polynucleotides are useful for detecting antibodies in
 CC a biological sample. Ligands that bind to the proteins and antibodies

CC directed against the proteins are useful for detecting the proteins
CC and for detecting HBsAg and HBeAg (precursor proteins). They
CC are also useful for preparing a vaccine or medicament for treating
CC HBV infections. The present sequence is encoded by the genome of
CC HBV genotype G strain FRI.
XX

SO Sequence 399 AA;

Query Match 90.4%; Score 1743; DB 22; Length 399;
Best Local Similarity 90.4%; Pred. No. 3.7e-128;
Matches 309; Conservative 13; Mismatches 20; Indels 0; Gaps 0;

```

GAGAGGAGLGPFPFHGGLGWSPOAGILFTLPANPPASTNRSGRQPTPLSPPLRNTHPQ 65
58 GVGAGTGPFTPHGGLGWSPOAGILFTLPANPPASTNRSGRQPTPLSPPLRNTHPQ 117
66 AMQNSTTFHQTLODPVRGLYFPAGSSSGTVNVPPTVPSISIFSRIDPALNMENI 125
118 AMQNSTAFHQAIDPNKVGILYFPAGSSSGTVNVPPTVPSISIFSRIDPALNMENI 177
126 TSGFLGPLVLVLAQGFPLRLITIPQSLDSWMTSLNPLGCTVCLGONSOSPSTNSHSPS 185
178 TSGFLGPLVLVLAQGFPLRLITIPQSLDSWMTSLNPLGCTVCLGONSOSPSTNSHSPS 237
186 CPPTCPGRMWCRLRRFIIFLLCLIFLLVLDYOGMLPVCPLIPGSSSTSTGPKCTC 245
238 CPPTCPGRMWCRLRRFIIFLLCLIFLLVLDYOGMLPVCPLIPGSSSTSTGPKCTC 297
246 TTPAGTSMYPSCCCTKPSDGNCTCIPSSMAFGKELMNASRPSWLSILVFPVQMFV 305
298 TTPAGTSMYPSCCCTKPSDGNCTCIPSSMAFGKELMNASRPSWLSILVFPVQMFV 357
306 GUSPTVWLSVIWMWYMGPSLSILSPFLPLPIFFCLMWYI 347
358 GUSPTVWLSVIWMWYMGPSLSILSPFLPLPIFFCLMWYI 399

```

RESULT 12

AAP80416
ID AAP80416 standard; protein; 383 AA.

AAAP80416;

12-SEP-1990 (first entry)

Sequence of adp type hepatitis B virus (HBV) surface antigen
(HBsAg) L protein (M protein, S protein).

XX Egg white lysozyme; hepatitis B virus (HBV) surface antigen (HBsAg);
KW Saccharomyces cerevisiae AH22R/PGLD Lp39-RCT; L protein; M protein;
KM Saccharomyces cerevisiae L11P39-RCT; Saccharomyces cerevisiae Lp31-RCT;
S protein.
XX

OS Hepatitis B virus.

XX EP288198-A.

XX 26-OCT-1988.

XX 13-APR-1988; 88EP-0303297.

XX 12-OCT-1987; 87JP-0256885.

XX (TAKE) TAKEDA CHEMICAL IND KK.

XX Fujisawa Y, Imai S, Miyazaki T;

XX WPI; 1988-301233/43.

XX N-PSDB; AAN80973.

XX Peptide (s) having hepatitis B surface antigenicity
XX used as vaccine for prevention of hepatitis B virus infection
XX and in diagnostic kits

XX Example; Fig 5; 31pp; English.

PS HBsAg L protein (P39) is an env protein. When L protein genes are
XX introduced into animal cells (CHO cells) they produce HBsAg particles
XX containing M and S proteins. The advantage is that peptides having
XX HBsAg antigenicity are excreted outside the cells and easily purified.
CC Claimed is a eukaryotic cell transformed with the rDNA. Pref.
CC the cell is a yeast cell, esp. S. cerevisiae AH22R-/PGLD Lp39-
CC RCT, L11P39-RCT or Lp31-RCT. Also claimed is recombinant DNA, which
CC is a DNA coding for a signal peptide which functions in a eukaryotic
CC cell, and is bound to the 5'-terminal of a DNA coding for a peptide
CC having HBsAg activity. The signal peptide may be a signal peptide of
CC egg white lysozyme. When trypsin-like protease-producing yeast is
CC utilized as the host, L protein and M protein may possibly be degraded
CC by the protease; therefore it is desirable that the genes are altered
CC so that the 48th arginine residue from the N-terminal of M protein or a
CC peptide containing the residue (preferably the 44-49 peptide) may be
CC deleted. Preferable genes include L protein gene (altered) coding for
CC the amino acid sequence 1-383 shown in AAN80973.

XX Sequence 383 AA;

Query Match 90.0%; Score 1736; DB 9; Length 383;
Best Local Similarity 90.6%; Pred. No. 1.2e-127;
Matches 310; Conservative 10; Mismatches 16; Indels 6; Gaps 1;

```

GAGAGGAGLGPFPFHGGLGWSPOAGILFTLPANPPASTNRSGRQPTPLSPPLRNTHPQ 65
48 GAGAGTGPFTPHGGLGWSPOAGILFTLPANPPASTNRSGRQPTPLSPPLRNTHPQ 107
66 AMQNSTTFHQTLODPVRGLYFPAGSSSGTVNVPPTVPSISIFSRIDPALNMENI 125
108 AMQNSTTFHQAIDPNKVGILYFPAGSSSGTVNVPPTVPSISIFSRIDPALNMENI 161
126 TSGFLGPLVLVLAQGFPLRLITIPQSLDSWMTSLNPLGCTVCLGONSOSPSTNSHSPS 185
162 TSGFLGPLVLVLAQGFPLRLITIPQSLDSWMTSLNPLGCTVCLGONSOSPSTNSHSPS 221
186 CPPTCPGRMWCRLRRFIIFLLCLIFLLVLDYOGMLPVCPLIPGSSSTSTGPKCTC 245
222 CPPTCPGRMWCRLRRFIIFLLCLIFLLVLDYOGMLPVCPLIPGSSSTSTGPKCTC 281
246 TTPAGTSMYPSCCCTKPSDGNCTCIPSSMAFGKELMNASRPSWLSILVFPVQMFV 305
282 TTPAGTSMYPSCCCTKPSDGNCTCIPSSMAFGKELMNASRPSWLSILVFPVQMFV 341
306 GUSPTVWLSVIWMWYMGPSLSILSPFLPLPIFFCLMWYI 347
342 GUSPTVWLSVIWMWYMGPSLSILSPFLPLPIFFCLMWYI 383

```

RESULT 13

AAR55287
ID AAR55287 standard; protein; 389 AA.

AAAR55287;

02-JAN-1995 (first entry)

Deduced sequence of HBV surface antigen LS.

XX Surface antigen; recombinant replicable vaccinia virus; hepatitis B
XX prevention; therapy; epitope; hepatitis B virus.
XX

OS Hepatitis B virus.

XX WO9412617-A.

XX 09-JUN-1994.

XX 24-NOV-1993; 93WO-US11474.

Query	Subject	Score	DB	Length	Gaps
25-NOV-1992;	92US-0982211.				
PA	(ITBI-) INT BIOTECHNOLOGY LAB INC.				
PI	Bernstine EG, Lewis T, Okeefe RM, Souw PTS;				
XX	WPI; 1994-200247/24.				
DR	N-FSDB; AA067970.				
PT	Prevention and treatment of hepatitis - using recombinant				
PT	replicable vaccinia viruses cony. hepatitis B virus surface and				
XX	core antigen nucleotide sequences				
XX	Example; Fig 22; 252bp; English.				
CC	The HBV genome (subtype adw) is contained in plasmid pAM6. The				
CC	3' end of the ORF coding for the three surface Ags was manipulated				
CC	to generate PRO-02 which contains an intact ORF capable of encoding				
CC	LS, MS and S Ags. The DNA sequences encoding S, MS and LS were sub-				
CC	cloned and sequenced. The sequences of the LS are AA067970/RS5287;				
CC	the sequences of MS are AA067961/RS5281 and The sequences of S are				
CC	AA067962/RS5282. Changes in nucleotide and deduced AA sequences				
CC	from the published sequence are believed to result from variability				
XX	within the adw subtype.				
XX	Sequence 389 AA;				
XX	Query Match 89.5%; Score 1727; DB 15; Length 389;				
XX	Best Local Similarity 89.2%; Pred. No. 6.3e-127;				
XX	Matches 305; Conservative 14; Mismatches 23; Indels 0; Gaps 0				
QY	6 GAGAGFAGCTPPHGHGLGMSPOAGILFETLPANPAPASTNRSGROPTLSPPLANTHQ 65				
DB	48 GVGAFGPGFTPPHGGVLGMSPOAGILITVSTTPASTNRSGROPTLSPPLRNSHPQ 107				
QY	66 AANMNSTTHQTLQDPKRVHGLTFPAGSSSGTVNPVPTVSPISISISIGDPALMNI 125				
DB	108 AANMNSTTHQALQDPKRVHGLTFPAGSSSGTVNPAPNASHISSISANTGDPVNMNI 167				
QY	136 TSGFLGPLVLVADGFLLTRILTIPOSLSNMTSLNFIQCTIVCLGQNSQSPISNHSPTS 185				
DB	168 TSGFLGPLVLVADGFLLTRILTIPOSLSNMTSLNFIQSGSPVCLGQNSQSPISNHSPTS 227				
QY	186 CPPTCPGRMNCLEPFIIFLFIILLCLIFLVLVDYQGMVPCPILIPGSSSTSTGCPRTC 245				
DB	228 CPPIPCGRMNCLEPFIIFLFIILLCLIFLVLVDYQGMVPCPILIPGTTSTGCPRTC 287				
QY	246 TTPAGTSTVYPCCCCTKPSDNCCTCLIPSSNAFGKFLWENASAPSWLSLVPVQMFV 305				
DB	288 TTPAGTSTVYPCCCCTKPLDNCCTCLIPSSNAFAKYLWENASVFPWSLILVPVQMFV 347				
QY	306 GLSPFTWLSVIMWMYWGPSLVSILSPFLPLPIFCCLMVIY 347				
DB	348 GLSPFTWLSVIMWMYWGPSLVSIVSPFLPLPIFCCLMVIY 389				
XX	RESULT 14				
XX	AAVS4045				
XX	AAVS4045 standard; Protein; 400 AA.				
XX	AAVS4045;				
XX	27-MAR-2000 (first entry)				
XX	Amino acid sequence of a HBV large surface antigen protein.				
XX	HBV; HBV surface antigen-'S'-133 Oon strain (Met to Thr);				
XX	DNA polymerase; large surface antigen; core protein;				
XX	transactivating X protein; hepatitis vaccine; HBV infection;				
XX	hepatocellular carcinoma.				
XX	Hepatitis B virus.				

[illegible]

Db 359 GLSPTVWLSVIMMMWYGRSLYNILSPFLPLPIFFCLMVYI 400

RESULT 15

AA10851
ID AA10851 standard; Protein; 388 AA.

XX AA10851;

XX 29-APR-1991 (first entry)

XX Modified hepatitis B virus large protein (1).

XX Hepatitis B virus; large surface protein; L protein; myristylation;

XX vaccines.

XX Hepatitis B virus.

XX Key Location/Qualifiers

XX Region 1..107

XX Region 108..162

XX Region 163..388

XX Region 163..388

XX EP414374-A.

XX 27-FEB-1991.

XX 19-JUL-1990; 90EP-0307900.

XX 03-AUG-1989; 89US-0389184.

XX 25-JUL-1989; 89US-0385342.

XX (SMIK) SMITHKLINE BIOLOGIC.

XX Comberbach M, Harford N, Gabazon T, Rutgers A, Voet P,

XX Jacobs E, Hollenberg CP, Janowicz ZA, Merckelberg AJ;

XX WPI, 1991-055585/09.

XX N-PSDB; AAQ10678.

XX Modified hepatitis B virus large surface protein - has amino acid

XX sequence encoding L protein used in vaccine for treating or

XX preventing hepatitis B without effects

XX Disclosure; Page 4-7; 79pp; English.

XX The modified L protein has the Gly 13 amino acid in the

XX Cc presl region of the wild-type sequence deleted. The deletion

XX results in the synthesis of a non-myristylated L protein.

XX (The first amino of the ad sequence are not given).

XX The modified product may be used alone or in a composite particle

XX comprising at least two polypeptides corresp. to all or part of a

XX protein having the biological activity of one of the hepatitis

XX Cc B surface antigens. The composite particles are useful for

XX Cc preparing improved hepatitis B vaccines or for treating of HBV.

XX See also AAQ10677-79.

XX Sequence 388 AA;

XX Query Match 88.8%; Score 1713; DB 12; Length 388;

XX Best Local Similarity 88.3%; Pred. No. 7.8e-126;

XX Matches 302; Conservative 15; Mismatches 25; Indels 0; Gaps 0;

QY 6 GAGAGLGLFTPPHGGGLGMSPOAGIILETLANPPASTNROSGROPTPLSRNTHPQ 65

DB 47 GVGAGFGPGLTPPHGGGLGMSPOAGIILETLVSTTPPASTNROSGROPTPLSRDSDHPQ 106

QY 66 AMQNSTFHQTLQDPRVRGLYFPAGSSSGGVNFPPTVSPISISIFGRIGDPALNMENI 125

Db 107 AMQNSTFHQTLQDPRVRGLYFPAGSSSGGVNFPPTVSPISISISARTGDPVTNMENI 166

QY 126 TSGFLGFLVLQAGFFLLTRITITPQSLDSWTSINFLGTTVCUGNSQSPTSNHSPTS 185

Db 167 TSGFLGFLVLQAGFFLLTRITITPQSLDSWTSINFLGSPVCUGNSQSPTSNHSPTS 226

QY 186 CPPTCPGYRWMCLRRFIFLFLVLLCLFPLVLDYQMLPVCPLIPGSSSTTSTGFCRTC 245

Db 227 CPPTCPGYRWMCLRRFIFLFLVLLCLFPLVLDYQMLPVCPLIPGSSSTTSTGFCRTC 286

QY 246 TTPAGTSMYPSCCTKPSDGNCTCIPSSWAFAFKPLMEWASAPSMLSLVPVQWVF 305

Db 287 TTPAGTSMYPSCCTKPSDGNCTCIPSSWAFAFKPLMEWASAPSMLSLVPVQWVF 346

QY 306 GLSPTVWLSVIMMMWYGRSLYNILSPFLPLPIFFCLMVYI 347

Db 347 GLSPTVWLSVIMMMWYGRSLYNILSPFLPLPIFFCLMVYI 388

Search completed: February 3, 2003, 09:23:50
Job time : 57.6167 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OW protein - protein search, using sw model

Run on: February 3, 2003, 09:21:12 ; Search time 28.8415 Seconds

(Without alignments)
1535.988 Million cell updates/sec

Title: US-09-890-752A-2

Sequence: 1 MPUSIFRSRIGPTVQASKL.....RRRSQSPRRRSQSRPQC 215

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP_PROTOZOA:*
12: SP_VIRUS:*
13: SP_VIRIDIA:*
14: SP_UNCLASSIFIED:*
15: SP_VIRUS:*
16: SP_BACTERIAP:*
17: SP_ARCHAEP:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1048.5	91.3	212	12	011884 hepatitis b
2	1040.5	90.6	212	12	089656 hepatitis b
3	1038.5	90.5	212	12	068016 hepatitis b
4	1037.5	90.4	212	12	067876 hepatitis b
5	1037.5	90.4	212	12	068020 hepatitis b
6	1037.5	90.4	212	12	089597 hepatitis b
7	1035.5	90.2	212	12	068025 hepatitis b
8	1035.5	90.2	212	12	068068 hepatitis b
9	1034.5	90.1	212	12	067884 hepatitis b
10	1034.5	90.1	212	12	068032 hepatitis b
11	1034.5	89.9	212	12	068070 hepatitis b
12	1032.5	89.9	212	12	091577 hepatitis b
13	1032.5	89.9	212	12	091717 hepatitis b
14	1032.5	89.9	212	12	0910K2 hepatitis b
15	1031.5	89.9	212	12	067872 hepatitis b
16	1031.5	89.9	212	12	068051 hepatitis b

17	1030.5	89.8	212	12	091T01	091c01 hepatitis b
18	1030.5	89.8	212	12	091Sx4	091s44 hepatitis b
19	1030.5	89.8	212	12	067980	067980 hepatitis b
20	1030.5	89.8	212	12	068012	068012 hepatitis b
21	1030.5	89.8	212	12	068014	068014 hepatitis b
22	1029.5	89.7	212	12	091S25	091s25 hepatitis b
23	1029.5	89.7	212	12	091S21	091s21 hepatitis b
24	1029.5	89.7	212	12	091S74	091s74 hepatitis b
25	1029.5	89.7	212	12	091U03	091u03 hepatitis b
26	1029.5	89.7	212	12	0910N8	0910n8 hepatitis b
27	1029.5	89.7	212	12	068010	068010 hepatitis b
28	1029.5	89.7	212	12	068075	068075 hepatitis b
29	1028.5	89.6	212	12	096845	096845 hepatitis b
30	1028.5	89.6	212	12	090M48	090m48 hepatitis b
31	1028.5	89.6	212	12	091705	091705 hepatitis b
32	1028.5	89.6	212	12	068030	068030 hepatitis b
33	1028.5	89.6	212	12	089446	089446 hepatitis b
34	1027.5	89.5	212	12	091Sx8	091s8 hepatitis b
35	1027.5	89.5	212	12	0910L4	0910l4 hepatitis b
36	1026.5	89.4	212	12	091T03	091t03 hepatitis b
37	1026.5	89.4	212	12	091S26	091s26 hepatitis b
38	1026.5	89.4	212	12	091S20	091s20 hepatitis b
39	1026.5	89.4	212	12	091S78	091s78 hepatitis b
40	1026.5	89.4	212	12	091S76	091s76 hepatitis b
41	1026.5	89.4	212	12	091S75	091s75 hepatitis b
42	1026.5	89.4	212	12	091S73	091s73 hepatitis b
43	1026.5	89.4	212	12	091707	091707 hepatitis b
44	1026.5	89.4	212	12	0910L8	0910l8 hepatitis b
45	1025.5	89.3	212	12	091T04	091t04 hepatitis b

ALIGNMENTS

RESULT 1
ID 011884 PRELIMINARY; PRT; 212 AA.
AC 011884;
DT 01-JUL-1997 (TEMBUREL 04, Created)
DT 01-JUL-1997 (TEMBUREL 04, Last sequence update)
DT 01-DEC-2001 (TEMBUREL 19, Last annotation update)
DR Core antigen.
OS Hepatitis B virus.
OC Viruses; Retroviruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.
RA Rao B.S., Casey J.L., Rinaldo J.S., Korba B.E.,
RT "Complete nucleotide sequence of a molecular clone of hepatitis B
RT virus from the 2.2.15 cell line."
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U95551; AAB55971.1;
DR InterPro: IPR002006; Hepatitis_core.
DR Pfam: PF00906; Hepatitis_core; 1.
SQ SEQUENCE 212 AA; 24360 MW; D301689EBB05A1B8 CRC64;

Query Match 91.3%; Score 1048.5; DB 12; Length 212;
Best local Similarity 97.5%; Pred. No. 2.5e-98;
Matches 198; Conservative 0; Mismatches 0; Indels 5; Gaps 1;
QY 13 PTYQASKLCIGMWMNDIDPYKFGATVLLSFLPSDFPVSVDLIDVYSALYREALSP 72
DB 15 PTYQASKLCIGMWMNDIDPYKFGATVLLSFLPSDFPVSVDLIDVYSALYREALSP 74
QY 73 BHOSPHHTLRQAIIICWGBTLMTATVGNLDDPFRGDASGLVVSYYNTMGKFRQL 132
DB 75 BHOSPHHTLRQAIIICWGBTLMTATVGNLDDPFRGDASGLVVSYYNTMGKFRQL 129
QY 133 LWFHISCLTFGRRETVIEYLVSGVWIRTPPAYRPNNAPILSTLPTTVRRGRSPRRRT 192
DB 130 LWFHISCLTFGRRETVIEYLVSGVWIRTPPAYRPNNAPILSTLPTTVRRGRSPRRRT 189
QY 193 PSPRRRSQSPRRRSQSRPQC 215

RN [1]
 RP SEQUENCE FROM N.A.
 RA Lai M.E., Mazzoleni A.P., Balestrieri A., Melis A., Porru A.;
 RT "Sequence analysis of HBV genomes isolated from patients with HBeAg
 negative chronic liver disease."
 RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
 DR EMBL; X65258; CAA46354.1; -
 DR InterPro; IPR002006; Hepatitis_core.
 DR Pfam; PF00906; Hepatitis_core; 1.
 SQ SEQUENCE 212 AA; 24336 MW; 6E91CC7D068EB573 CRC64;

Query Match 90.4%; Score 1037.5; DB 12; Length 212;
 Best Local Similarity 96.6%; Pred. No. 3.3e-97;
 Matches 196; Conservative 1; Mismatches 1; Indels 5; Gaps 1;

Qy 13 PTVQASKLCGLMGMDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALESP 72
 Db 15 PTVQASKLCGLMGMDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALESP 74
 Qy 73 EHCSPHHTALROAILCNGELMTLATWGVNLDPFRGDASRDLYVSYVNTNGLKFRQL 132
 Db 75 EHCSPHHTALROAILCNGELMTLATWGVNLDPFRGDASRDLYVSYVNTNGLKFRQL 129
 Qy 133 LMFHISCLTFGRRTVLEYLVSFGVWIRTPPAYRPNAPILSTLPTTVVRRGRSPRRRT 192
 Db 130 LMFHISCLTFGRRTVLEYLVSFGVWIRTPPAYRPNAPILSTLPTTVVRRGRSPRRRT 189
 Qy 193 PSPRRRSQSPRRRSQSPRRSQC 215
 Db 190 PSPRRRSQSPRRRSQSPRRSQC 212

RESULT 5
 Q68020 PRELIMINARY; PRT; 212 AA.

AC Q68020;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Pre-C/core protein.
 GN PRE-C/CORE.
 OS Hepatitis B virus.
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
 OX NCBI_TaxID=10407;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lai M.E., Mazzoleni A.P., Porru A., Balestrieri A.;
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL; X85291; CAA59611.1; -
 DR InterPro; IPR002006; Hepatitis_core.
 DR Pfam; PF00906; Hepatitis_core; 1.
 SQ SEQUENCE 212 AA; 24336 MW; 63006A4EBB04A1B8 CRC64;

Query Match 90.4%; Score 1037.5; DB 12; Length 212;
 Best Local Similarity 96.6%; Pred. No. 3.3e-97;
 Matches 196; Conservative 1; Mismatches 1; Indels 5; Gaps 1;

Qy 13 PTVQASKLCGLMGMDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALESP 72
 Db 15 PTVQASKLCGLMGMDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALESP 74
 Qy 73 EHCSPHHTALROAILCNGELMTLATWGVNLDPFRGDASRDLYVSYVNTNGLKFRQL 132
 Db 75 EHCSPHHTALROAILCNGELMTLATWGVNLDPFRGDASRDLYVSYVNTNGLKFRQL 129
 Qy 133 LMFHISCLTFGRRTVLEYLVSFGVWIRTPPAYRPNAPILSTLPTTVVRRGRSPRRRT 192
 Db 130 LMFHISCLTFGRRTVLEYLVSFGVWIRTPPAYRPNAPILSTLPTTVVRRGRSPRRRT 189
 Qy 193 PSPRRRSQSPRRRSQSPRRSQC 215
 Db 190 PSPRRRSQSPRRRSQSPRRSQC 212

RESULT 6

Q68597 PRELIMINARY; PRT; 212 AA.

AC Q68597;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE HBcAg (PRE-CORE protein) (PRECORE/core protein).
 OS Hepatitis B virus (subtype ayw), and
 OS Hepatitis B virus.
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
 OX NCBI_TaxID=10418, 10407;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Hepatitis B virus (subtype ayw); STRAIN=SUB-TYPE AYW;
 RX MEDLINE=94079539; PubMed=8257295;
 RA Pfeister-Adams S., Schlayer M.J., Peters T., Hettler F., Gerok W.,
 Raenack J.;
 RT "Sequence analysis of hepatitis B virus DNA in immunologically
 negative infection."
 RL Arch. Virol. 133:385-396(1993).

RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Hepatitis B virus (subtype ayw); STRAIN=AYW;
 RA Karayannis P.;
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Hepatitis B virus (subtype ayw); STRAIN=AYW;
 RA Karayannis P.;
 RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.

RN [4]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Hepatitis B virus; STRAIN=143-PAM G;
 RA Zampino R., Lobello S., Chiaramonte M., Venturi-Pasini C., Dumble U.,
 Thuerz M., Karayannis P.;
 RT "Intra-familial transmission of Hepatitis B virus in Italy."
 RT Phylogenetic sequence analysis and amino acid variation of the core
 RT gene."
 RL J. Hepatol. 0:0-0(2002).

RN [5]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Hepatitis B virus; STRAIN=GAM1119F29, AND GAM1821F217;
 RA Dumble U., Mendy M., Karayannis P.;
 RT "Prevalence of HBV core promoter/precore/core mutations in Gambian
 RT chronic carriers."
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; X72702; CAA51257.1; -
 DR EMBL; X80925; CAA56887.1; -
 DR EMBL; AF419525; AAL15953.1; -
 DR EMBL; AF350205; AAK57324.1; -
 DR EMBL; AF350205; AAK57324.1; -
 DR InterPro; IPR002006; Hepatitis_core.
 DR Pfam; PF00906; Hepatitis_core; 1.
 FT CHAIN 30 212 CORE PROTEIN.
 SQ SEQUENCE 212 AA; 24336 MW; 1861B13B8B047AC8 CRC64;

Query Match 90.4%; Score 1037.5; DB 12; Length 212;
 Best Local Similarity 96.6%; Pred. No. 3.3e-97;
 Matches 196; Conservative 1; Mismatches 1; Indels 5; Gaps 1;

Qy 13 PTVQASKLCGLMGMDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALESP 72
 Db 15 PTVQASKLCGLMGMDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALESP 74
 Qy 73 EHCSPHHTALROAILCNGELMTLATWGVNLDPFRGDASRDLYVSYVNTNGLKFRQL 132
 Db 75 EHCSPHHTALROAILCNGELMTLATWGVNLDPFRGDASRDLYVSYVNTNGLKFRQL 129
 Qy 133 LMFHISCLTFGRRTVLEYLVSFGVWIRTPPAYRPNAPILSTLPTTVVRRGRSPRRRT 192
 Db 130 LMFHISCLTFGRRTVLEYLVSFGVWIRTPPAYRPNAPILSTLPTTVVRRGRSPRRRT 189

QY 193 PPSRRRSQSPRRRSQSPREPOC 215
 DB 190 PPSRRRSQSPRRRSQSPREPOC 212

RESULT 7

ID Q68025 PRELIMINARY; PRT; 212 AA.
 AC Q68025;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Pre-C/core protein.
 GN Pre-C/CORE.
 OS Hepatitis B virus.
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
 OX NCBI_TaxID=10407;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lai M.E., Mazzoleni A.P., Portu A., Balestrieri A.;
 RL Submitted (MAR-1995) to the EMBL/Genbank/DBJ databases.
 DR EMBL; X85293; CA59616.1; -
 DR InterPro; IPR002006; Hepatitis core.
 DR Pfam; PF00906; Hepatitis core; 1
 SQ SEQUENCE 212 AA; 24363 MW; 6940D04E90A410C3 CRC64;

Query Match 90.2%; Score 1035.5; DB 12; Length 212;
 Best Local Similarity 96.6%; Pred. No. 5.3e-97;
 Matches 196; Conservative 0; Mismatches 2; Indels 5; Gaps 1;

QY 13 PTVOASKLCGLMGMDIDPYKEFGATVLLSFLPSDFPVSVDLDTASALYREALSP 72
 DB 15 PTVOASKLCGLMGMDIDPYKEFGATVLLSFLPSDFPVSVDLDTASALYREALSP 74
 QY 73 EHCSPHHTALROAILCWEGLMTLATVGVNLEDP-----ASRDVSVSYNTNMGLKFRQL 132
 DB 75 EHCSPHHTALROAILCWEGLMTLATVGVNLEDP-----ASRDVSVSYNTNMGLKFRQL 129
 QY 133 LMFHISCLTFRGRTVIEYLVSGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRT 192
 DB 130 LMFHISCLTFRGRTVIEYLVSGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRT 189
 QY 193 PPSRRRSQSPRRRSQSPREPOC 215
 DB 190 PPSRRRSQSPRRRSQSPREPOC 212

RESULT 8

ID Q68068 PRELIMINARY; PRT; 212 AA.
 AC Q68068;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Pre-C/core protein.
 GN Pre-C/CORE.
 OS Hepatitis B virus.
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
 OX NCBI_TaxID=10407;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lai M.E., Mazzoleni A.P., Portu A., Balestrieri A.;
 RL Submitted (MAR-1995) to the EMBL/Genbank/DBJ databases.
 DR EMBL; X85315; CA59664.1; -
 DR InterPro; IPR002006; Hepatitis core.
 DR Pfam; PF00906; Hepatitis core; 1
 SQ SEQUENCE 212 AA; 24366 MW; CEABDD4E8B1B5E8 CRC64;

Query Match 90.2%; Score 1035.5; DB 12; Length 212;
 Best Local Similarity 96.6%; Pred. No. 5.3e-97;
 Matches 196; Conservative 0; Mismatches 2; Indels 5; Gaps 1;

QY 13 PTVOASKLCGLMGMDIDPYKEFGATVLLSFLPSDFPVSVDLDTASALYREALSP 72
 DB 15 PTVOASKLCGLMGMDIDPYKEFGATVLLSFLPSDFPVSVDLDTASALYREALSP 74
 QY 73 EHCSPHHTALROAILCWEGLMTLATVGVNLEDP-----ASRDVSVSYNTNMGLKFRQL 132
 DB 75 EHCSPHHTALROAILCWEGLMTLATVGVNLEDP-----ASRDVSVSYNTNMGLKFRQL 129
 QY 133 LMFHISCLTFRGRTVIEYLVSGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRT 192
 DB 130 LMFHISCLTFRGRTVIEYLVSGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRT 189
 QY 193 PPSRRRSQSPRRRSQSPREPOC 215
 DB 190 PPSRRRSQSPRRRSQSPREPOC 212

RESULT 9

ID Q67984 PRELIMINARY; PRT; 212 AA.
 AC Q67984;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Pre-C/core protein.
 GN Pre-C/CORE.
 OS Hepatitis B virus.
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
 OX NCBI_TaxID=10407;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lai M.E., Mazzoleni A.P., Portu A., Balestrieri A.;
 RL Submitted (MAR-1995) to the EMBL/Genbank/DBJ databases.
 DR EMBL; X85256; CA59519.1; -
 DR InterPro; IPR002006; Hepatitis core.
 DR Pfam; PF00906; Hepatitis core; 1
 SQ SEQUENCE 212 AA; 24322 MW; 629E0B298B13DD01 CRC64;

Query Match 90.1%; Score 1034.5; DB 12; Length 212;
 Best Local Similarity 96.6%; Pred. No. 6.7e-97;
 Matches 196; Conservative 0; Mismatches 2; Indels 5; Gaps 1;

QY 13 PTVOASKLCGLMGMDIDPYKEFGATVLLSFLPSDFPVSVDLDTASALYREALSP 72
 DB 15 PTVOASKLCGLMGMDIDPYKEFGATVLLSFLPSDFPVSVDLDTASALYREALSP 74
 QY 73 EHCSPHHTALROAILCWEGLMTLATVGVNLEDP-----ASRDVSVSYNTNMGLKFRQL 132
 DB 75 EHCSPHHTALROAILCWEGLMTLATVGVNLEDP-----ASRDVSVSYNTNMGLKFRQL 129
 QY 133 LMFHISCLTFRGRTVIEYLVSGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRT 192
 DB 130 LMFHISCLTFRGRTVIEYLVSGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRT 189
 QY 193 PPSRRRSQSPRRRSQSPREPOC 215
 DB 190 PPSRRRSQSPRRRSQSPREPOC 212

RESULT 10

ID Q68032 PRELIMINARY; PRT; 212 AA.
 AC Q68032;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Pre-C/core protein.
 GN Pre-C/CORE.
 OS Hepatitis B virus.
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
 OX NCBI_TaxID=10407;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Lai M.E., Mazzeoli A.P., Portu A., Balestrieri A.;
 RU Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: X85296; CA59622.1; -
 DT InterPro: IPR002006; Hepatitis_core.
 DR Pfam: PF00906; Hepatitis_core; 1.
 SQ SEQUENCE 212 AA; 24308 MW; E28972C78B1604FF CRC64;

Query Match 90.1%; Score 1034.5; DB 12; Length 212;
 Best Local Similarity 96.1%; Pred. No. 6,7e-97;
 Matches 195; Conservative 2; Mismatches 1; Indels 5; Gaps 1;

QY 13 PTVQASKLCGLMGMDIDPYKEGATVETLSFLPSDFPSPVRDLDTASALYREALSP 72
 DB 15 PTVQASKLCGLMGMDIDPYKEGATVETLSFLPSDFPSPVRDLDTASALYREALSP 74
 QY 73 EHCSPHHTALROAILCWEIMTLATWGVNLEDPEFRGDSRDVLVSYVNTNMGKRFOL 132
 DB 75 EHCSPHHTALROAILCWEIMTLATWGVNLEDPEFRGDSRDVLVSYVNTNMGKRFOL 129
 QY 133 LMFHISCLTGFRETVLYLVSFGVWIRTPPAYRPNPAPILSTLPTTVVRRGRSPRRRT 192
 DB 130 LMFHISCLTGFRETVLYLVSFGVWIRTPPAYRPNPAPILSTLPTTVVRRGRSPRRRT 189
 QY 193 PSPRRRSQSPRRRSQSPRRSQC 215
 DB 190 PSPRRRSQSPRRRSQSPRRSQC 212

RESULT 11

ID 068070 PRELIMINARY; PRT; 212 AA.
 AC 068070;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 GN Pre-C/core protein.
 OS Hepatitis B virus.
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
 NC NCB1_TaxID=10407;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lai M.E., Mazzeoli A.P., Portu A., Balestrieri A.;
 RU Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: X85316; CA59666.1; -
 DT InterPro: IPR002006; Hepatitis_core.
 DR Pfam: PF00906; Hepatitis_core; 1.
 SQ SEQUENCE 212 AA; 24376 MW; D30173662996A1B8 CRC64;

Query Match 90.1%; Score 1034.5; DB 12; Length 212;
 Best Local Similarity 96.6%; Pred. No. 6,7e-97;
 Matches 196; Conservative 0; Mismatches 2; Indels 5; Gaps 1;

QY 13 PTVQASKLCGLMGMDIDPYKEGATVETLSFLPSDFPSPVRDLDTASALYREALSP 72
 DB 15 PTVQASKLCGLMGMDIDPYKEGATVETLSFLPSDFPSPVRDLDTASALYREALSP 74
 QY 73 EHCSPHHTALROAILCWEIMTLATWGVNLEDPEFRGDSRDVLVSYVNTNMGKRFOL 132
 DB 75 EHCSPHHTALROAILCWEIMTLATWGVNLEDPEFRGDSRDVLVSYVNTNMGKRFOL 129
 QY 133 LMFHISCLTGFRETVLYLVSFGVWIRTPPAYRPNPAPILSTLPTTVVRRGRSPRRRT 192
 DB 130 LMFHISCLTGFRETVLYLVSFGVWIRTPPAYRPNPAPILSTLPTTVVRRGRSPRRRT 189
 QY 193 PSPRRRSQSPRRRSQSPRRSQC 215
 DB 190 PSPRRRSQSPRRRSQSPRRSQC 212

RESULT 12
 0915Y7 PRELIMINARY; PRT; 212 AA.

AC 0915Y7;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DS Precore/core protein.
 OS Hepatitis B virus.
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
 NC NCB1_TaxID=10407;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GAM1211;
 RA Dumps U., Mendy M., Karayannis P.;
 RT "Prevalence of HBV core promoter/precure/core mutations in Gambian
 RT chronic carriers";
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF350155; AA657272.1; -
 DT InterPro: IPR002006; Hepatitis_core.
 DR Pfam: PF00906; Hepatitis_core; 1.
 SQ SEQUENCE 212 AA; 24363 MW; 27E96FD069BC35A CRC64;

Query Match 89.9%; Score 1032.5; DB 12; Length 212;
 Best Local Similarity 96.1%; Pred. No. 1.1e-96;
 Matches 195; Conservative 1; Mismatches 2; Indels 5; Gaps 1;

QY 13 PTVQASKLCGLMGMDIDPYKEGATVETLSFLPSDFPSPVRDLDTASALYREALSP 72
 DB 15 PTVQASKLCGLMGMDIDPYKEGATVETLSFLPSDFPSPVRDLDTASALYREALSP 74
 QY 73 EHCSPHHTALROAILCWEIMTLATWGVNLEDPEFRGDSRDVLVSYVNTNMGKRFOL 132
 DB 75 EHCSPHHTALROAILCWEIMTLATWGVNLEDPEFRGDSRDVLVSYVNTNMGKRFOL 129
 QY 133 LMFHISCLTGFRETVLYLVSFGVWIRTPPAYRPNPAPILSTLPTTVVRRGRSPRRRT 192
 DB 130 LMFHISCLTGFRETVLYLVSFGVWIRTPPAYRPNPAPILSTLPTTVVRRGRSPRRRT 189
 QY 193 PSPRRRSQSPRRRSQSPRRSQC 215
 DB 190 PSPRRRSQSPRRRSQSPRRSQC 212

RESULT 13

ID 091717 PRELIMINARY; PRT; 212 AA.
 AC 091717;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 GN Pre-core protein.
 OS Hepatitis B virus.
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
 NC NCB1_TaxID=10407;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=144-PAM G;
 RA Zampino R., Lobello S., Chiaramonte M., Venturi-Pasini C., Dumps U.,
 RA Thurez M., Karayannis P.;
 RT "Intra-familial transmission of Hepatitis B virus in Italy:
 RT Phylogenetic sequence analysis and amino acid variation of the core
 RT gene";
 RL J. Hepatol. 0:0-0(2002).
 DR EMBL: AF419526; AA15954.1; -
 DT InterPro: IPR002006; Hepatitis_core.
 DR Pfam: PF00906; Hepatitis_core; 1.
 FT CHAIN 30 212 CORE PROTEIN.
 SQ SEQUENCE 212 AA; 24421 MW; B76FE275B047AD1 CRC64;

Query Match 89.9%; Score 1032.5; DB 12; Length 212;
 Best Local Similarity 96.1%; Pred. No. 1.1e-96;
 Matches 195; Conservative 1; Mismatches 2; Indels 5; Gaps 1;

QY 13 PTVQASKLCGLMGMDIDPYKEGATVETLSFLPSDFPSPVRDLDTASALYREALSP 72

```

Db      15 PTVQASKLCGLWGMQMDIDPYKEFGATVELLSFLPSDFPSPVRDLDTASALYDALESP 74
Oy      73 EHCSPHRTALFOAILCWCGLMTLATWGVNLDEPFEGDASRDVSVYVNTNMGKFRQL 132
Db      75 EHCSPHRTALFOAILCWCGLMTLATWGVNLDEP-----ASRDVSVYVNTNMGKFRQL 129
Oy      133 LMFHISCLTGFRETVIELVSPFGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRT 192
Db      130 LMFHISCLTGFRETVIELVSPFGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRT 189
Oy      193 PSPRRRRSQSPRRRRSQSPRRPQC 215
Db      190 PSPRRRRSQSPRRRRSQSPRRSQC 212

RESULT 14
Oy10K2 PRELIMINARY; PRT; 212 AA.
ID      O910K2;
AC      01-DEC-2001 (TREMblrel. 19, Created)
DT      01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT      01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE      Precore/core protein.
OS      Hepatitis B virus.
OC      Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
NC      NCBI_TaxId=10407;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Dumps U., Wendy M., Karayannis P.;
RT      "Prevalence of HBV core promoter/precure/core mutations in Gambian
RT      chronic carriers."
RT      Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.
DR      EMBL; AF350118; AAK57335.1; -
DR      EMBL; AF350099; AAK57216.1; -
DR      EMBL; AF350100; AAK57217.1; -
DR      EMBL; AF350102; AAK57219.1; -
DR      EMBL; AF350103; AAK57220.1; -
DR      EMBL; AF350106; AAK57223.1; -
DR      EMBL; AF350108; AAK57225.1; -
DR      EMBL; AF350116; AAK57233.1; -
DR      EMBL; AF350118; AAK57235.1; -
DR      EMBL; AF350119; AAK57236.1; -
DR      EMBL; AF350120; AAK57237.1; -
DR      EMBL; AF350121; AAK57238.1; -
DR      EMBL; AF350122; AAK57239.1; -
DR      EMBL; AF350124; AAK57241.1; -
DR      EMBL; AF350125; AAK57242.1; -
DR      EMBL; AF350126; AAK57243.1; -
DR      EMBL; AF350128; AAK57245.1; -
DR      EMBL; AF350130; AAK57247.1; -
DR      EMBL; AF350132; AAK57249.1; -
DR      EMBL; AF350133; AAK57250.1; -
DR      EMBL; AF350138; AAK57255.1; -
DR      EMBL; AF350140; AAK57257.1; -
DR      EMBL; AF350141; AAK57258.1; -
DR      EMBL; AF350143; AAK57260.1; -
DR      EMBL; AF350147; AAK57264.1; -
DR      EMBL; AF350148; AAK57265.1; -
DR      EMBL; AF350149; AAK57266.1; -
DR      EMBL; AF350150; AAK57267.1; -
DR      EMBL; AF350151; AAK57268.1; -
DR      EMBL; AF350152; AAK57269.1; -
DR      EMBL; AF350156; AAK57273.1; -
DR      EMBL; AF350157; AAK57274.1; -
DR      EMBL; AF350162; AAK57279.1; -
DR      EMBL; AF350163; AAK57280.1; -
DR      EMBL; AF350165; AAK57282.1; -
DR      EMBL; AF350169; AAK57286.1; -
DR      EMBL; AF350170; AAK57287.1; -
DR      EMBL; AF350171; AAK57288.1; -
DR      EMBL; AF350172; AAK57289.1; -
DR      EMBL; AF350174; AAK57291.1; -

```

```

DR      EMBL; AF350178; AAK57295.1; -
DR      EMBL; AF350181; AAK57298.1; -
DR      EMBL; AF350185; AAK57302.1; -
DR      EMBL; AF350186; AAK57303.1; -
DR      EMBL; AF350189; AAK57306.1; -
DR      EMBL; AF350190; AAK57307.1; -
DR      EMBL; AF350192; AAK57309.1; -
DR      EMBL; AF350193; AAK57310.1; -
DR      EMBL; AF350194; AAK57311.1; -
DR      EMBL; AF350195; AAK57312.1; -
DR      EMBL; AF350196; AAK57313.1; -
DR      EMBL; AF350197; AAK57314.1; -
DR      EMBL; AF350198; AAK57315.1; -
DR      EMBL; AF350199; AAK57316.1; -
DR      EMBL; AF350200; AAK57317.1; -
DR      EMBL; AF350201; AAK57318.1; -
DR      EMBL; AF350210; AAK57327.1; -
DR      EMBL; AF350212; AAK57329.1; -
DR      EMBL; AF350213; AAK57330.1; -
DR      EMBL; AF350214; AAK57331.1; -
DR      EMBL; AF350216; AAK57333.1; -
DR      EMBL; AF350217; AAK57334.1; -
DR      EMBL; AF350217; AAK57334.1; -
DR      InterPro: IPR002006; Hepatitis_core.
DR      Pfam: PF00906; Hepatitis_core; 1.
SQ      SEQUENCE 212 AA; 24349 MW; 1A61B13EBB10C583 CRC64;

```

```

Query Match      89.9%; Score 1032.5; DB 12; Length 212;
Best Local Similarity 96.1%; Pred. No. 1,1e-96;
Matches 195; Conservative 1; Mismatches 2; Indels 5; Gaps 1;

```

```

Oy      13 PTVQASKLCGLWGMQMDIDPYKEFGATVELLSFLPSDFPSPVRDLDTASALYDALESP 72
Db      15 PTVQASKLCGLWGMQMDIDPYKEFGATVELLSFLPSDFPSPVRDLDTASALYDALESP 74
Oy      73 EHCSPHRTALFOAILCWCGLMTLATWGVNLDEPFEGDASRDVSVYVNTNMGKFRQL 132
Db      75 EHCSPHRTALFOAILCWCGLMTLATWGVNLDEP-----ASRDVSVYVNTNMGKFRQL 129
Oy      133 LMFHISCLTGFRETVIELVSPFGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRT 192
Db      130 LMFHISCLTGFRETVIELVSPFGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRT 189
Oy      193 PSPRRRRSQSPRRRRSQSPRRPQC 215
Db      190 PSPRRRRSQSPRRRRSQSPRRSQC 212

RESULT 15
Oy10K2 PRELIMINARY; PRT; 212 AA.
ID      O67872;
AC      067872;
DT      01-NOV-1996 (TREMblrel. 01, Created)
DT      01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT      01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE      Pre C/C ORF.
OS      Hepatitis B virus.
OC      Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
NC      NCBI_TaxId=10407;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Lai M.R., Mazzeletti A.P., Balestrieri A., Mellis A., Porru A.;
RT      "Sequence analysis of HBV genomes isolated from patients with HBsAg
RT      negative chronic liver disease."
RT      Submitted (MAR-1992) to the EMBL/Genbank/DBJ databases.
DR      EMBL; X65257; CAA46350.1; -
DR      InterPro: IPR002006; Hepatitis_core.
DR      Pfam: PF00906; Hepatitis_core; 1.
SQ      SEQUENCE 212 AA; 24348 MW; 1B61AB0DD21B49B1 CRC64;

```

```

Query Match      89.9%; Score 1031.5; DB 12; Length 212;
Best Local Similarity 96.1%; Pred. No. 1,1e-96;
Matches 195; Conservative 1; Mismatches 2; Indels 5; Gaps 1;

```

Qy 13 PTQASRLCJGMLGMDIDPYKRGATVELLSPLSPDPSPVRDLDTASALYREALESP 72
Db 15 PTQASRLCJGMLGMDIDPYKRGATVELLSPLSPDPSPVRDLDTASALYREALESP 74
Qy 73 EHCSPHHTALROAILCWEGLMTLATWVGNTLBDPEFRGDASRDLYVSYVNTNGLKFRQL 132
Db 75 EHCSPHHTALROAILCWEGLMTLATWVGNTLBDPEFRGDASRDLYVSYVNTNGLKFRQL 129
Qy 133 LMFHISCLITGRETIVBYLVSPGVWIRTPPAYRPPNAPILSTLPETTIVVRRGRSPRRRT 192
Db 130 LMFHISCLITGRETIVBYLVSPGVWIRTPPAYRPPNAPILSTLPETTIVVRRGRSPRRRT 189
Qy 193 PSPRRRSOSPPRRRSOSREPOC 215
Db 190 PSPRRRSOSPPRRRSOSRESQC 212

Search completed: February 3, 2003, 09:25:44
Job time : 30.1748 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd

OM protein - protein search, using SW model

Run on: February 3, 2003, 09:16:26 ; Search time 7.86585 Seconds
(without alignments)
113.686 Million cell updates/sec

```

Title: US-09-090-752A-2
Perfect score: 1148
Sequence: 1 MPLSLTFSRIGDPTVQASKL.....RRRRSQSPRRRRSQSNRPQC 215

```

scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

```
Searched:      112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters:  112892
```

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

Database : SwissProt_40:4

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	992.5	86.5	214	1	CORA_HBPV9	P17099 hepatitis i
2	988	86.1	211	1	CORA_HBPV4	P24023 hepatitis i
3	985.5	85.8	212	1	CORA_HBPV7	O05495 hepatitis i
4	951.5	82.9	183	1	CORA_HBPV1	P03146 hepatitis i
5	932.5	81.2	183	1	CORA_HBPV2	P03147 hepatitis i
6	919.5	80.1	183	1	CORA_HBPV3	P17391 hepatitis i
7	919.5	80.1	183	1	CORA_HBPV0	P17392 hepatitis i
8	919.5	80.1	185	1	CORA_HBPV4	P03149 hepatitis i
9	918.5	80.0	183	1	CORA_HBPV4	P03150 hepatitis i
10	912.5	79.5	185	1	CORA_HBPV2	P28178 hepatitis i
11	911.5	79.4	185	1	CORA_HBPV2	P03148 hepatitis i
12	899.5	78.4	183	1	CORA_HBPV3	P03149 hepatitis i
13	757.5	66.0	217	1	CORA_HBPV3	P03152 ground squa
14	671	58.4	188	1	CORA_HBV1	P03152 woodchuck i
15	655.5	57.1	187	1	CORA_HBV8	P06433 woodchuck i
16	175	15.2	305	1	CORA_HBV8	P13845 heron hepat
17	173	15.1	305	1	CORA_HBPDC	P30027 duck hepat
18	171	14.9	305	1	CORA_HBPDC	P11190 duck hepat
19	168	14.6	305	1	CORA_HBPV0	P17191 duck hepat
20	166	14.5	305	1	CORA_HBPV0	P03154 duck hepat
21	111	9.7	196	1	SFR2_CAEEL	O09511 caenorhabd
22	94.5	8.2	484	1	SFR8_HUMAN	O05519 homo sapien
23	90.5	7.9	208	1	YSX2_CAEEL	Q10021 caenorhabd
24	86	7.5	951	1	SFR8_HUMAN	Q10287 homo sapien
25	84	7.3	219	1	NOL3_HUMAN	O06935 homo sapien
26	84	7.3	498	1	VE2_HPV08	P06422 human papil
27	84	7.3	503	1	VE2_HPV21	P50767 human papil
28	80.5	7.0	329	1	SRA4_CAEEL	O09206 caenorhabd
29	80.5	7.0	318	1	HRD_FUGU9	P51112 fugu rubrip
30	80	7.0	832	1	AND3_HUMAN	P57078 homo sapien
31	80	7.0	867	1	VL96_IRV1	P22856 tipula irid
32	79.5	6.9	303	1	SFR1_NAHT	O22315 arabiszopsis
33	79.5	6.9	891	1	MAY1_SCHCO	P37936 schizophysil

34	79	6.9	483	1	VE2_HPV14	P26793	human papill
35	79	6.9	279	1	ED2_HUMAN	O95071	homo sapien
36	78	6.8	164	1	SPF3_HUMAN	P23152	homo sapien
37	78	6.8	1092	1	SOMA_DROME	P21297	drosophila
38	78	6.8	2426	1	SOM_HUMAN	P18563	homo sapien
39	77.5	6.8	252	1	VPHE_NPVAC	P24728	autographa
40	77.5	6.8	493	1	VE2_HPV19	P26766	human papill
41	77.5	6.8	673	1	LOX5_HUMAN	P09917	homo sapien
42	77.5	6.8	2404	1	SOM_MOUSE	O99x47	mus musculi
43	77	6.7	77	1	PR22_SEPOF	P80002	sepia offic
44	76	6.6	1286	1	CTRO_HUMAN	O14578	homo sapien
45	75.5	6.6	802	1	BDB2_ACERY	O82860	acetobacter

ALIGNMENTS

```

RESULT 1
COR_A HPBV9          STANDARD;          PRT;          214 AA.
ID   CORA_HPBV9
AC   P17059;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   Core antigen.
GN   C.
OS   Hepatitis B virus (subtype adv / strain 991).
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_Taxid:10410;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Koechel H.G., Schueler A., Lottmann S., Thomsen R.;
RL   Submitted (FEB-1990) to the EMBL/GenBank/DBJ databases.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation
CC   at the European Bioinformatics Institute. There are no restrictions on
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for
CC   entities requires a license agreement (See http://www.isb.ch/announce/
CC   or send an email to license@isb-sib.ch).
DR   EMBL, X51970; CAA36232.1; -.
DR   PIR, S10381; NKVLKS.
DR   InterPro, IPR002006; Hepatitis_core.
DR   Pfam, PF00906; Hepatitis_core.1.
KW   Core protein; Repeat.
KM   REPEAT
FT   REPEAT          193      200
FT   FT              201      208
SQ   SEQUENCE      214 AA; 24722 MW; 2D668333BC5AFB8C CRC64;
Query Match          86.5%; Score 992.5; DB 1; Length 214;
Best Local Similarity 93.2%; Pred. NO. 8.3e-82;
Matches 191; Conservative 2; Mismatches 5; Indels 7; Gaps 2;
QY   13 PTVQASKICUGMIMGMDIDPYKEFGATVELLSFPLSPDFPFSVRDLDTASALYREALESP 72
Db   15 PTVQASKICUGMIMGMDIDPYKEFGATVELLSFPLSPDFPFSVRDLDTASALYREALESP 74
QY   73 EHGGPHHTALROALLCMGELMTLATWVGNIENEDBFRGDASRDLYSVYVNTNMGIKFRQL 132
Db   75 EHGGPHHTALROALLCMGELMTLATWVGNNIEDP-----ASRDLYVNVVNTNMGIKIRQL 129
QY   133 LWFHISCLTGRETVIEYLVSGFWIRTPPAYRPPNAPILSTLPETVVR--RGRSPR 190
Db   130 LWFHISYLTFGRETVIEYLVSGFWINTPPAYRPPNAPILSTLPETVVRKRDGRSPR 189
QY   191 RTBSPRRRSOSPPRRRSOSRPPC 215
Db   190 RTBSPRRRSOSPPRRRSOSRPPC 214

```

CORA_HPBVA STANDARD; PRT; 211 AA.

AC P24023;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Core antigen.

GN Hepatitis B virus (strain alpha).
 OS Viruses; Retrovird viruses; Hepadnaviridae; Orthohepadnavirus.
 OC NCBI_TaxID=10411;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90266476; PubMed=2345966;
 RA Tong S., Li J., Vilevick L., Trepo C.;
 RT "Active hepatitis B virus replication in the presence of anti-HBe is
 associated with viral variants containing an inactive pre-C region.";
 RL Virology 176:596-603 (1990).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC
 CC EMBL; M32138; -; NOT_ANNOTATED_CDS.
 DR PIR; A34773; NKLAL.
 DR InterPro; IPR002006; Hepatitis_core.
 DR Pfam; PF00906; Hepatitis_core; 1.
 KW Core protein; Repeat.
 FT REPEAT 183 190
 FT REPEAT 198 206
 SQ SEQUENCE 211 AA; 24208 MW; B774AC72B55C75AB CRC64;

Query Match 86.1%; Score 988; DB 1; Length 211;
 Best Local Similarity 93.1%; Pred. No. 2,1e-81;
 Matches 189; Conservative 4; Mismatches 4; Indels 6; Gaps 2;

QY 13 PTVQASKLCGLMGMDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSP 72
 DB 15 PTVQASKLCGLMGMDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSP 73

QY 73 EHCSPHTALROAILCWSGLMTLATWGVNLEDEPFGRDASRDVLVSVYVNTMGLKFRQL 132
 DB 74 EHCSPHTALROAILCWSGLMTLATWGVNLEDEPFGRDASRDVLVSVYVNTMGLKFRQL 128

QY 133 LMFHISCLTGFRETVIELVSPGWIRTPPAYRPPNAPILSTLPETVVRRGRSPRRRT 192
 DB 129 LMFHISCLTGFRETVIELVSPGWIRTPPAYRPPNAPILSTLPETVVRRGRSPRRRT 188

QY 193 PSPRRRSQSPRRRSQSPRRSPOC 215
 DB 189 PSPRRRSQSPRRRSQSPRRSPOC 211

RESULT 3
 CORA_HPBVT STANDARD; PRT; 212 AA.
 AC 005495;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Core antigen.
 GN C.
 OS Hepatitis B virus (subtype adw4 / strain Brazil / isolate w4B).
 OS Viruses; Retrovird viruses; Hepadnaviridae; Orthohepadnavirus.
 OC NCBI_TaxID=45410;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93346970; PubMed=834535;
 RA Nefmann H., Schaefer S., Yoshida C.F.T., Gaspar A.M.C., Repp R.,

RA Gerlich W.H.;
 RT "Identification of a new hepatitis B virus (HBV) genotype from Brazil
 that expresses HBV surface antigen subtype adw4.";
 RL J. Gen. Virol. 74:1627-1632 (1993).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC
 CC EMBL; X69798; CAA9453.1;
 DR InterPro; IPR002006; Hepatitis_core.
 DR Pfam; PF00906; Hepatitis_core; 1.
 KW Core protein; Repeat.
 FT DOMAIN 178 204 ARG-RICH.
 FT DOMAIN 184 203 3 X 5 AA REPEATS OF S-P-R-R-R.
 FT REPEAT 184 188 1.
 FT REPEAT 191 195 2.
 FT REPEAT 199 203 3.
 SQ SEQUENCE 212 AA; 24234 MW; F832610DB7C36FD2 CRC64;

Query Match 85.8%; Score 985.5; DB 1; Length 212;
 Best Local Similarity 91.1%; Pred. No. 3.5e-81;
 Matches 185; Conservative 6; Mismatches 5; Indels 5; Gaps 1;

QY 13 PTVQASKLCGLMGMDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSP 72
 DB 15 PTVQASKLCGLMGMDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSP 74

QY 73 EHCSPHTALROAILCWSGLMTLATWGVNLEDEPFGRDASRDVLVSVYVNTMGLKFRQL 132
 DB 75 EHCSPHTALROAILCWSGLMTLATWGVNLEDEPFGRDASRDVLVSVYVNTMGLKFRQL 129

QY 133 LMFHISCLTGFRETVIELVSPGWIRTPPAYRPPNAPILSTLPETVVRRGRSPRRRT 192
 DB 130 LMFHISCLTGFRETVIELVSPGWIRTPPAYRPPNAPILSTLPETVVRRGRSPRRRT 189

QY 193 PSPRRRSQSPRRRSQSPRRSPOC 215
 DB 190 PSPRRRSQSPRRRSQSPRRSPOC 212

RESULT 4
 CORA_HPBVT STANDARD; PRT; 183 AA.
 AC P03146;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 01-APR-1990 (Rel. 14, Last annotation update)
 DE Core antigen.
 GN C.
 OS Hepatitis B virus (subtype ayw).
 OS Viruses; Retrovird viruses; Hepadnaviridae; Orthohepadnavirus.
 OC NCBI_TaxID=10418;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=81012051; PubMed=399327;
 RA Galibert F., Mandart B., Filoussi F., Toulais P., Charney P.;
 RT "Nucleotide sequence of the hepatitis B virus genome (subtype ayw)
 cloned in B. coli.";
 RL Nature 281:646-650 (1979).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: V01460; CAA24706.1; ALT_INIT.
DR EMBL: X02496; -; NOT_ANNOTATED_CDS.
DR PIR: A03711; NRYLAH.
DR PIR: A03712; NRYLAH.
DR InterPro: IPR002006; Hepatitis_core.
DR Pfam: PF00906; Hepatitis_core; 1.
KM Core protein; Repeat.
FT REPEAT 162 169
FT REPEAT 170 177
FT CONFLICT 33 33 T -> N (IN REF. 2).
FT CONFLICT 80 80 A -> I (IN REF. 2).
SQ SEQUENCE 183 AA; 21116 MW; E0D9D9763F24E958 CRC64;

Query Match
Best Local Similarity 82.9%; Score 951.5; DB 1; Length 183;
Matches 182; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

QY 28 MDIDPYKFGATVETLSFLPSDFPSVVDLDTASALYRBALSPHSCSPHHTALROAIL 87
DB 1 MDIDPYKFGATVETLSFLPSDFPSVVDLDTASALYRBALSPHSCSPHHTALROAIL 60

QY 88 CWGELMTLATWGVNLEDPFRGDASRDLVVSYNTNMGKFRQLMFWHISCLTFGRRTV 147
DB 61 CWGELMTLATWGVNLEDPFRGDASRDLVVSYNTNMGKFRQLMFWHISCLTFGRRTV 115

QY 148 IEYVSGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTPSPRRRSQSPRRR 207
DB 116 IEYVSGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTPSPRRRSQSPRRR 175

QY 208 SQSREPOC 215
DB 176 SQSRESOC 183

RESULT 5
CORR_HPBVZ STANDARD; PRT; 183 AA.
ID CORR_HPBVZ
AC P03147;
DT 21-JUN-1986 (Rel. 01, Created)
DT 21-JUN-1986 (Rel. 01, Last sequence update)
DT 01-OCT-1993 (Rel. 12, Last annotation update)
DE Core antigen.
GN C.
OS Hepatitis B virus (subtype adw).
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
CX NCBI_TaxID=10413;
RX SEQUENCE FROM N.A.
RA MEDLINE=8101215; PubMed=39329;
RA Paek M., Goto T., Gilbert W., Zink B., Schaller H., McKay P.,
RA Lebedev G., Murray K.,
RA "Hepatitis B virus genes and their expression in E. coli."
RT Nature 282:575-579 (1979).
RU
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: J03202; AAA45486.1; -;
DR EMBL: A08967; CAA00816.1; -;
DR PIR: B93217; NKYLA2.
DR InterPro: IPR002006; Hepatitis_core.

DR Pfam: PF00906; Hepatitis_core; 1.
KM Core protein; Repeat.
FT REPEAT 162 169
FT REPEAT 170 177
SQ SEQUENCE 183 AA; 21042 MW; 545ED0B5527F26C CRC64;

Query Match
Best Local Similarity 81.2%; Score 932.5; DB 1; Length 183;
Matches 176; Conservative 5; Mismatches 2; Indels 5; Gaps 1;

QY 28 MDIDPYKFGATVETLSFLPSDFPSVVDLDTASALYRBALSPHSCSPHHTALROAIL 87
DB 1 MDIDPYKFGATVETLSFLPSDFPSVVDLDTASALYRBALSPHSCSPHHTALROAIL 60

QY 88 CWGELMTLATWGVNLEDPFRGDASRDLVVSYNTNMGKFRQLMFWHISCLTFGRRTV 147
DB 61 CWGELMTLATWGVNLEDPFRGDASRDLVVSYNTNMGKFRQLMFWHISCLTFGRRTV 115

QY 148 IEYVSGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTPSPRRRSQSPRRR 207
DB 116 IEYVSGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTPSPRRRSQSPRRR 175

QY 208 SQSREPOC 215
DB 176 SQSRESOC 183

RESULT 6
CORR_HPBVZ STANDARD; PRT; 183 AA.
ID CORR_HPBVZ
AC P17391;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE Core antigen.
OS Hepatitis B virus (subtype adw / strain Japan/pUDW233).
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
CX NCBI_TaxID=10413;
RX SEQUENCE FROM N.A.
RA MEDLINE=89010694; PubMed=3171552;
RA Okamoto H., Tsuda F., Sakugawa H., Saetsoewigunjo R.I., Imai M.,
RA Miyakawa Y., Mayumi M.;
RA "Typing hepatitis B virus by homology in nucleotide sequence:
RT comparison of surface antigen subtypes";
RT J. Gen. Virol. 69:2575-2583 (1988).
RU
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: D00329; -; NOT_ANNOTATED_CDS.
DR PIR: A28925; NKYLA1.
DR InterPro: IPR002006; Hepatitis_core.
DR Pfam: PF00906; Hepatitis_core; 1.
KM Core protein; Repeat.
FT REPEAT 162 169
FT REPEAT 170 177
SQ SEQUENCE 183 AA; 21224 MW; 9FDD6B5F5AF5E160 CRC64;

Query Match
Best Local Similarity 80.1%; Score 919.5; DB 1; Length 183;
Matches 175; Conservative 3; Mismatches 5; Indels 5; Gaps 1;

QY 28 MDIDPYKFGATVETLSFLPSDFPSVVDLDTASALYRBALSPHSCSPHHTALROAIL 87
DB 1 MDIDPYKFGATVETLSFLPSDFPSVVDLDTASALYRBALSPHSCSPHHTALROAIL 60

QY 88 CWGELMTLATWGVNLEDPFRGDASRDLVVSYNTNMGKFRQLMFWHISCLTFGRRTV 147

```

Db 61 CMGELMTLATVGNLLEDP-----ASRDLYVNVYNTMGLKIRQLMFWHISCLTFGRET 115
Qy 148 IEYLVSGVWIRTPPAYRPPNAPILSTLPTTVVRRRRSPRRRSPRRRSOSPPRRR 207
Db 116 LEYLVSGVWIRTPPAYRPPNAPILSTLPTTVVRRRRSPRRRTPSPRRRSOSPPRRR 175
Qy 208 SOSREPOC 215
Db 176 SOSRESOC 183

RESULT 7
CORR_HPBV0
ID CORR_HPBV0 STANDARD, PRT: 183 AA.
AC P17392,
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE Core antigen.
OS Hepatitis B virus (subtype adr / strain Okinawa/PDWM282).
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
CC NCBI_TaxId=10415;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89010694; PubMed=3171552;
RA Okamoto H., Tenda F., Sakugawa H., Saitosocwignjo R.I., Imai M.,
RA Miyakawa Y., Mayumi M.;
RT "Typing hepatitis B virus by homology in nucleotide sequence:
RT comparison of surface antigen subtypes."
RL J. Gen. Virol. 69:2575-2583(1988).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
CC EMBL: D00330; -; NOT_ANNOTATED_CDS.
DR PIR: B28925; NRYLJ2.
DR InterPro: IPR002006; Hepatitis_core.
DR Pfam: PF00906; Hepatitis_core; 1.
KW Core protein; Repeat.
FT REPEAT 162 169
FT REPEAT 170 177
FT REPEAT 172 179
SQ SEQUENCE 183 AA; 21081 MW; 66BDB2633122930C CRC64;

Query Match 80.1%; Score 919.5; DB 1; Length 183;
Best Local Similarity 93.1%; Pred. No. 2.3e-75;
Matches 175; Conservative 3; Mismatches 5; Indels 5; Gaps 1;

Qy 28 MDIDPYKFGATVLLSPIDSPSVRDLDPTASALYREALSPHCSPHNTALRQAIL 87
Db 1 MDIDPYKFGATVLLSPIDSPSVRDLDPTASALYREALSPHCSPHNTALRQAIL 60
Qy 88 CMGELMTLATVGNLLEDPERFGDASRDLYVSYNTMGLKFRQLMFWHISCLTFGRET 147
Db 61 CMGELMTLATVGNLLEDP-----ASRDLYVSYNTMGLKIRQLMFWHISCLTFGRET 115
Qy 148 IEYLVSGVWIRTPPAYRPPNAPILSTLPTTVVRRRRSPRRRTPSPRRRSOSPPRRR 207
Db 116 LEYLVSGVWIRTPPAYRPPNAPILSTLPTTVVRRRRSPRRRTPSPRRRSOSPPRRR 175
Qy 208 SOSREPOC 215
Db 176 SOSRESOC 183

RESULT 8
CORR_HPBV0
ID CORR_HPBV0 STANDARD, PRT: 185 AA.

```

```

AC P03149;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-OCT-1989 (Rel. 12, Last annotation update)
DE Core antigen.
GN C.
OS Hepatitis B virus (subtype adr).
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
CC NCBI_TaxId=106821;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83168919; PubMed=6300776;
RA Ono Y., Ono H., Sasada R., Igashahi K., Sugino Y., Nishioaka K.;
RT "The complete nucleotide sequences of the cloned hepatitis B virus
RT DNA; subtype adr and adr."
RL Nucleic Acids Res. 11:1747-1757(1983).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
CC EMBL: V00866; -; NOT_ANNOTATED_CDS.
DR PIR: C93460; NRYLAE.
DR InterPro: IPR002006; Hepatitis_core.
DR Pfam: PF00906; Hepatitis_core; 1.
KW Core protein; Repeat.
FT REPEAT 164 171
FT REPEAT 172 179
FT REPEAT 174 179
SQ SEQUENCE 185 AA; 21394 MW; B86A90D541BA70F9 CRC64;

Query Match 80.1%; Score 919.5; DB 1; Length 185;
Best Local Similarity 93.2%; Pred. No. 2.4e-75;
Matches 177; Conservative 3; Mismatches 3; Indels 7; Gaps 2;

Qy 28 MDIDPYKFGATVLLSPIDSPSVRDLDPTASALYREALSPHCSPHNTALRQAIL 87
Db 1 MDIDPYKFGATVLLSPIDSPSVRDLDPTASALYREALSPHCSPHNTALRQAIL 60
Qy 88 CMGELMTLATVGNLLEDPERFGDASRDLYVSYNTMGLKFRQLMFWHISCLTFGRET 147
Db 61 CMGELMTLATVGNLLEDP-----ASRDLYVSYNTMGLKIRQLMFWHISCLTFGRET 115
Qy 148 IEYLVSGVWIRTPPAYRPPNAPILSTLPTTVVRRRRSPRRRTPSPRRRSOSPPRRR 205
Db 116 LEYLVSGVWIRTPPAYRPPNAPILSTLPTTVVRRRRSPRRRTPSPRRRSOSPPRRR 175
Qy 206 RRSOSREPOC 215
Db 176 RRSOSRESOC 185

RESULT 9
CORR_HPBV4
ID CORR_HPBV4 STANDARD, PRT: 183 AA.
AC P03150; P03151;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE Core antigen.
GN C.
OS Hepatitis B virus (subtype adr), and
OS Hepatitis B virus (subtype adr / strain Indonesia/p1DM420).
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
CC NCBI_TaxId=10409, 106820, 10412;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ADR;
RX MEDLINE=83168919; PubMed=6300776;

```

RA Ono Y., Onda H., Sasada R., Igarashi K., Sugino Y., Nishika K.;
 RT "The complete nucleotide sequences of the cloned hepatitis B virus
 RT DNA; subtype adr and adw.";
 RL Nucleic Acids Res. 11:1747-1757(1983).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ADRA;
 RX MEDLINE=83246570; PubMed=6306594;
 RA Fujiyama A., Miyano H., Nozaki C., Yoneyama T., Ohmoto N.,
 RA Matsubara K.;
 RT "Cloning and structural analyses of hepatitis B virus DNAs, subtype
 RT adr.";
 RL Nucleic Acids Res. 11:4601-4610(1983).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ADW;
 RX MEDLINE=89010694; PubMed=3171552;
 RA Okamoto H., Teuda F., Sakugawa H., Saitosowegimjo R.I., Imai M.,
 RA Miyakawa Y., Mayumi M.;
 RT "Typing hepatitis B virus by homology in nucleotide sequence:
 RT comparison of surface antigen subtypes.";
 RL J. Gen. Virol. 69:2575-2583(1988).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-ch.ch).
 CC -----
 DR EMBL; V00667; -; NOT ANNOTATED_CDS.
 DR EMBL; X01587; CA25745.1; -;
 DR EMBL; D00331; -; NOT ANNOTATED_CDS.
 DR PIR; A93480; NKVLA4.
 DR PIR; B93460; NKVLA4.
 DR PIR; C28925; NKVLA3.
 DR InterPro: IPR002006; Hepatitis core.
 DR Pfam: PF00906; Hepatitis_core; 1.
 DR Core protein; Repeat.
 FT REPEAT 162 169
 FT REPEAT 170 177
 SO SEQUENCE 183 AA; 21095 MW; ED2DAIDB07F8596D CRC64;
 Query Match 80.4%; Score 918.5; DB 1; Length 183;
 Best Local Similarity 92.6%; Pred. No. 2.9e-75;
 Matches 174; Conservative 4; Mismatches 5; Indels 5; Gaps 1;
 QY 28 MDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSPHSCPHHTALRQAIL 87
 DB 1 MDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSPHSCPHHTALRQAIL 60
 QY 88 CWGELMTLATVGNLDEPERGASRLVSVYNTMGLKFRQLMFWHISCLTFGRRTV 147
 DB 61 CWGELMTLATVGNLDEPERGASRLVSVYNTMGLKFRQLMFWHISCLTFGRRTV 115
 QY 148 IEYVSPGFWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRTSPRRRSQSPRRR 207
 DB 116 IEYVSPGFWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRTSPRRRSQSPRRR 175
 QY 208 SOSREPOC 215
 DB 176 SOSRESQC 183
 RESULT 10
 CORA_HPBVF STANDARD; PRT; 195 AA.
 ID CORA_HPBVF
 AC P29178;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 01-DEC-1992 (Rel. 24, Last annotation update)
 DE Core antigen.

GN C.
 OS Hepatitis B virus (subtype adw2 variant sf).
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
 OX NCBI_TaxId=31515;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90169850; PubMed=2307406;
 RA Bhat R.A., Ulrich P.P., Vyas G.N.;
 RT "Molecular characterization of a new variant of hepatitis B virus in
 RT a persistently infected homosexual man.";
 RL Hepatology 11:271-276(1990).
 DR PIR; A93182; NKVLA3.
 DR InterPro: IPR002006; Hepatitis core.
 DR Pfam: PF00906; Hepatitis_core; 1.
 DR Core protein; Repeat.
 FT REPEAT 174 181
 FT REPEAT 182 189
 SO SEQUENCE 195 AA; 22461 MW; B2B16F879C7C87 CRC64;
 Query Match 79.5%; Score 912.5; DB 1; Length 195;
 Best Local Similarity 91.1%; Pred. No. 1.1e-74;
 Matches 174; Conservative 6; Mismatches 6; Indels 5; Gaps 1;
 QY 25 LMGMDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSPHSCPHHTALRQ 84
 DB 10 LMGMDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSPHSCPHHTALRQ 69
 QY 85 ALLCWGELMTLATVGNLDEPERGASRLVSVYNTMGLKFRQLMFWHISCLTFGR 144
 DB 70 ALLCWGELMTLATVGNLDEPERGASRLVSVYNTMGLKFRQLMFWHISCLTFGR 124
 QY 145 ETVIEYVSPGFWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRTSPRRRSQSPR 204
 DB 125 ETVIEYVSPGFWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRTSPRRRSQSPR 184
 QY 205 RRSQSPRPOC 215
 DB 185 RRSQSPRPOC 195
 RESULT 11
 CORA_HPBV2 STANDARD; PRT; 185 AA.
 ID CORA_HPBV2
 AC P03148;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Core antigen.
 GN C.
 OS Hepatitis B virus (subtype adw2).
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
 OX NCBI_TaxId=10408;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Valenzuela P., Quiroga M., Zaldívar J., Gray P., Rutter W.J.;
 RL (in) Field B.N., Jaenisch R., Fox C.F. (eds.);
 RL Animal virus genetics, pp.57-70, Academic Press, New York (1980).
 DR PIR; A94409; NKVLA3.
 DR InterPro: IPR002006; Hepatitis_core.
 DR Pfam: PF00906; Hepatitis_core; 1.
 DR Core protein; Repeat.
 FT REPEAT 164 171
 FT REPEAT 172 179
 SO SEQUENCE 185 AA; 21304 MW; 31F4DC38B507B19 CRC64;
 Query Match 79.4%; Score 911.5; DB 1; Length 185;
 Best Local Similarity 92.6%; Pred. No. 1.2e-74;
 Matches 176; Conservative 3; Mismatches 4; Indels 7; Gaps 2;
 QY 28 MDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSPHSCPHHTALRQAIL 87
 DB 1 MDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSPHSCPHHTALRQAIL 60

```

QY 88 CWGELMTLATVGVNLEDEPERGASRDVSVYNTNMGKLFKROLWPHISCLTGRETV 147
DB 61 CWGELMTLATVGVNLEDEP-----ASRDVSVYNTNMGKLFKROLWPHISCLTGRETV 115
QY 148 IEYLVSPGVWIRTPPAYRPPNAPILSTLPTETTVRR--RGRSPRRRTSPRRRSQSPRR 205
DB 116 IEYLVSPGVWIRTPPAYRPPNAPILSTLPTETTVRRRRGRSPRRRTSPRRRSQSPRR 175
QY 206 RRSQSPRRPQC 215
DB 176 RRSQSPRRSQC 185

RESULT 12
CORR_HPBVL STANDARD; PRT; 183 AA.
ID CORR_HPBVL
AC P12901;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-APR-1990 (Rel. 14, Last annotation update)
DE Core antigen.
GN C.
OS Hepatitis B virus (strain 1sh / chimpanzee isolate).
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
CC NCBI_TaxId=10414;
CC [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=88258473; PubMed=2838576;
RA Vaudin M., Wolstenholme A.J., Tsiangaye K.N., Zuckerman A.J.,
RA Harrison T.J.;
RT "The complete nucleotide sequence of the genome of a hepatitis B
RT virus isolated from a naturally infected chimpanzee."
RL J. Gen. Virol. 69:11383-11389(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D00220; BAA00157.1; -
DR PIR: A28865; NKTULP.
DR InterPro: IPR002006; Hepatitis_core.
DR Pfam: PF00906; Hepatitis_core; 1.
KW Core protein; Repeat.
FT REPEAT 162 169
FT REPEAT 170 177
SQ SEQUENCE 183 AA; 20999 MW; 923DCB94A33FC0B8 CRC64;

Query Match 78.4%; Score 899.5; DB 1; Length 183;
Best Local Similarity 91.5%; Pred. No. 1.4e-73;
Matches 172; Conservative 5; Mismatches 6; Indels 5; Gaps 1;

QY 28 MDIDPYKEGATVLLSFLSPDPSPVRLDLDPAALYREALESPHCSPHHTALROAIL 87
DB 1 MDIDPYKEGATVLLSFLSPDPSPVRLDLDPAALYREALESPHCSPHHTALROAIL 60
QY 88 CWGELMTLATVGVNLEDEPERGASRDVSVYNTNMGKLFKROLWPHISCLTGRETV 147
DB 61 CWGELMTLATVGVNLEDEP-----ASRDVSVYNTNMGKLFKROLWPHISCLTGRETV 115
QY 148 IEYLVSPGVWIRTPPAYRPPNAPILSTLPTETTVRRRRGRSPRRRTSPRRRSQSPRRR 207
DB 116 IEYLVSPGVWIRTPPAYRPPNAPILSTLPTETTVRRRRGRSPRRRTSPRRRSQSPRRR 175
QY 208 RRSQSPRRPQC 215
DB 176 RRSQSPRRSQC 183

RESULT 13

```

```

CORR_HPBGS STANDARD; PRT; 217 AA.
ID CORR_HPBGS
AC P03153;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-JUN-1990 (Rel. 13, Last annotation update)
DE Core antigen.
GN C.
OS Ground squirrel hepatitis virus (GSV).
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
CC NCBI_TaxId=10406;
CC [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=84267998; PubMed=6086950;
RA Seeger C., Ganem D., Varmus H.B.;
RT "Nucleotide sequence of an infectious molecularly cloned genome of
RT ground squirrel hepatitis virus."
RL J. Virol. 51:367-375(1984).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: K02715; AAA46755.1; -
DR PIR: A03715; NKTULS.
DR InterPro: IPR002006; Hepatitis_core.
DR Pfam: PF00906; Hepatitis_core; 1.
KW Core protein; Repeat.
FT REPEAT 196 203
FT REPEAT 204 211
SQ SEQUENCE 217 AA; 25189 MW; DF489467355SEC11A CRC64;

Query Match 66.0%; Score 757.5; DB 1; Length 217;
Best Local Similarity 65.6%; Pred. No. 8.7e-61;
Matches 143; Conservative 22; Mismatches 42; Indels 11; Gaps 2;

QY 3 LSSIFSRGDTTVQASKICLWLMGMDIDPYKEGATVLLSFLSPDPSPVRLDLPAS 62
DB 6 LCLVPAVCPCPTVQASKICLWLMGMDIDPYKEGSSQQLNLFPLDPPLNALVDITAA 65
QY 63 ALYREALESPHCSPHHTALROAILCWGELMTLATVGVNLEDEPERGASRDVSVYV 122
DB 66 ALYREALESPHCSPHHTALROALVCWELTRITMSENTE-----EEVRIIVDHVN 119
QY 123 TNMGKLFKROLWPHISCLTGRETVIEYLVSPGVWIRTPPAYRPPNAPILSTLPTETTVR 182
DB 120 TNMGKLFKROLWPHISCLTGFQHTVQVEFLVSPGVWIRTPPAYRPPNAPILSTLPTETTVIR 179
QY 183 RRG-----RSPRRRTSPRRRSQSPRRRSQSPRRPQC 215
DB 180 RRGSGRAASPPRRRTSPRRRSQSPRRRSQSPASNC 217

RESULT 14
CORR_HBV1 STANDARD; PRT; 188 AA.
ID CORR_HBV1
AC P03152;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Core antigen.
GN C.
OS Woodchuck hepatitis virus 1 (WHV 1),
OS Woodchuck hepatitis virus 7 (WHV 7),
OS Woodchuck hepatitis virus 59 (WHV 59), and
OS Woodchuck hepatitis virus 8 (infectious clone) (WHV 8).
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
CC NCBI_TaxId=10430, 10432, 10431, 10434;
CC [1]

```

```

RP SEQUENCE FROM N.A.
RC SPECIES=Woodchuck hepatitis virus 1;
RX MEDLINE=82216369; PubMed=7086958;
RA Gallibert F., Chen T.N., Mandart E.;
RT "Nucleotide sequence of a cloned woodchuck hepatitis virus genome:
RL J. Virol. 41:51-65(1982).
[2]
RN SEQUENCE FROM N.A.
RC SPECIES=Woodchuck hepatitis virus 7, and Woodchuck hepatitis virus 59;
RX MEDLINE=88101359; PubMed=3336938;
RA Cohen J.I., Miller R.H., Rosenblum B., Denniston K., Gerin J.L.,
RA Purcell R.H.;
RT "Sequence comparison of woodchuck hepatitis virus replicative forms
RL shows conservation of the genome.";
RN J. Virol. 162:12-20(1988).
[3]
RN SEQUENCE FROM N.A.
RC SPECIES=Woodchuck hepatitis virus 8 (infectious clone);
RX MEDLINE=89184524; PubMed=2928306;
RA Girones R., Cole P.J., Hornbuckle W.B., Tennant B.C., Gerin J.L.,
RA Purcell R.H., Miller R.H.;
RT "Complete nucleotide sequence of a molecular clone of woodchuck
hepatitis virus that is infectious in the natural host.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:1846-1849(1989).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J02442; AAA46761.1; -
DR EMBL; M18752; AAA46769.1; -
DR EMBL; M19183; AAA46765.1; -
DR EMBL; J04514; AAA46772.1; -
DR PIR; A03713; NKVLC.
DR PIR; C32397; NKVLC.
DR PIR; C32397; NKVLC.
DR PIR; C32397; NKVLC.
DR InterPro; IPR002006; Hepatitis_core.
DR Pfam; PF00906; Hepatitis_core.1.
DR Core protein; Repeat.
FT REPEAT 167 174
FT REPEAT 175 182
SQ SEQUENCE 188 AA; 15454D0A7B7CE42 CRC64;
Query Match 58.4%; Score 671; DB 1; Length 188;
Best Local Similarity 65.3%; Pred. No. 4e-53;
Matches 126; Conservative 21; Mismatches 36; Indels 10; Gaps 2;
QY 28 MDIDPKFGATVETLSPDPPSVRLDPTASALYREALSPHSCPHHTALROAIL 87
DB 1 MDIDPKFGSSYQLNPLPLDPPDLNALVDPTALYEEELTGRHSCPHHTAIRQALV 60
QY 88 CMGELMTLATVGVNLDEPFRGDASRDVLSYVNTMGKFKROLMPHISCLTFGRBTY 147
DB 61 CMDELTKLIAMSSNITSEGV-----RTIIVNVNDTGLKTRQSLMFHISCLTFGQHTV 115
QY 148 IEYLVSGVWIRTPPAYRPPNAPILSTLPTTVVRRG-----RSPRRTPSPRRRSOS 202
DB 116 QEFLVSGVWIRTPPAYRPPNAPILSTLPTTVVRRG-----RSPRRTPSPRRRSOS 202
QY 203 PRRRSOSRBPQC 215
DB 176 PRRRSOSPSANC 188

```

```

AC P06433;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 01-OCT-1989 (Rel. 12, Last annotation update)
DB Core antigen.
GN C.
OS Woodchuck hepatitis virus 8 (WHV 8).
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10433;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86062931; PubMed=3855246;
RA Kodama K., Ogasawara N., Yoshikawa H., Murakami S.;
RT "Nucleotide sequence of a cloned woodchuck hepatitis virus genome:
RL evolutionary relationship between hepadnaviruses.";
RN J. Virol. 56:978-986(1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M10882; AAA19185.1; -
DR PIR; A03714; NKVLC2.
DR InterPro; IPR002006; Hepatitis_core.
DR Pfam; PF00906; Hepatitis_core.1.
DR Core protein; Repeat.
FT REPEAT 166 173
FT REPEAT 174 181
SQ SEQUENCE 187 AA; 21579 MW; D4BC446FF7163165 CRC64;
Query Match 57.1%; Score 655.5; DB 1; Length 187;
Best Local Similarity 64.6%; Pred. No. 9.8e-52;
Matches 124; Conservative 21; Mismatches 38; Indels 9; Gaps 2;
QY 28 MDIDPKFGATVETLSPDPPSVRLDPTASALYREALSPHSCPHHTALROAIL 87
DB 1 MDIDPKFGSSYQLNPLPLDPPDLNALVDPTALYEEELTGRHSCPHHTAIRQALV 60
QY 88 CMGELMTLATVGVNLDEPFRGDASRDVLSYVNTMGKFKROLMPHISCLTFGRBTY 147
DB 61 CMDELTKLIAMSSNITSEGV-----RTIIVNVNDTGLKTRQSLMFHISCLTFGQHTV 115
QY 148 IEYLVSGVWIRTPPAYRPPNAPILSTLPTTVVRRG-----RSPRRTPSPRRRSOS 203
DB 116 QEFLVSGVWIRTPPAYRPPNAPILSTLPTTVVRRG-----RSPRRTPSPRRRSOS 203
QY 204 PRRRSOSRBPQC 215
DB 176 PRRRSOSPSANC 187

```

Search completed: February 3, 2003, 09:24:20
 Job time : 9.86585 secs

RESULT 15
 CORA_MHV8 STANDARD; PRT; 187 AA.
 ID CORA_MHV8

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 09:21:57 ; Search time 14.2334 Seconds
(without alignments)
1452.136 Million cell updates/sec

Title: US-09-890-752a-2

Perfect score: 1148

Sequence: 1 MFLSSIFSRIGDPTVQASKL.....RRRSQSPRRRRSQSREPOC 215

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 73:*

1: p1r1:.*
2: p1r2:.*
3: p1r3:.*
4: p1r4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1040.5	90.6	212	1 NKVLAH	e antigen precursor
2	1038.5	90.5	212	2 S53204	e antigen precursor
3	1037.5	90.4	212	2 S53211	e antigen precursor
4	1037.5	90.4	212	2 S53204	e antigen precursor
5	1037.5	90.4	212	2 S20750	e antigen precursor
6	1035.5	90.2	212	2 S53216	e antigen precursor
7	1035.5	90.2	212	2 S53212	e antigen precursor
8	1034.5	90.1	212	2 S53225	e antigen precursor
9	1034.5	90.1	212	2 S53274	e antigen precursor
10	1034.5	90.1	212	2 S53163	e antigen precursor
11	1031.5	89.9	212	2 S53211	e antigen precursor
12	1031.5	89.9	212	2 S20746	e antigen precursor
13	1030.5	89.8	212	2 S53129	e antigen precursor
14	1030.5	89.8	212	2 S53200	e antigen precursor
15	1030.5	89.8	212	2 S53202	e antigen precursor
16	1029.5	89.7	212	2 S53281	e antigen precursor
17	1029.5	89.7	212	2 S53198	e antigen precursor
18	1028.5	89.6	212	2 S53233	e antigen precursor
19	1028.5	89.6	212	2 S53227	e antigen precursor
20	1025.5	89.3	212	2 S53253	e antigen precursor
21	1023.5	89.2	212	1 NKVLBH	e antigen precursor
22	1022.5	89.1	212	2 S53240	e antigen precursor
23	1020.5	88.9	212	2 S53242	e antigen precursor
24	1016.5	88.5	212	2 S53257	e antigen precursor
25	1015.5	88.5	212	2 S53279	e antigen precursor
26	1014.5	88.4	212	2 S53229	e antigen precursor
27	1014.5	88.4	212	2 S53251	e antigen precursor
28	1012.5	88.2	212	2 S53255	e antigen precursor
29	1011.5	88.1	212	2 S53238	e antigen precursor

30	1011.5	88.1	214	2 S01405	core antigen - hep
31	1008.5	87.8	212	1 NKVLJ1	e antigen precursor
32	1008.5	87.8	212	1 NKVLJ2	e antigen precursor
33	1008.5	87.8	214	1 NKVLJ6	e antigen precursor
34	1007.5	87.8	212	1 NKVLJ4	e antigen precursor
35	1007.5	87.8	212	2 S53236	e antigen precursor
36	1007.5	87.8	212	2 S53157	e antigen precursor
37	1006.5	87.7	214	2 S47409	e antigen precursor
38	1002.5	87.3	214	2 S33686	e antigen precursor
39	1001.5	87.2	214	1 NKVLJ3	e antigen precursor
40	999.5	87.1	212	2 S53161	e antigen precursor
41	998.5	87.1	359	2 S04570	gene X/C fusion pr
42	992.5	86.5	214	1 NKVLK5	e antigen precursor
43	988	86.1	211	1 NKVLJ1	e antigen precursor
44	985.5	85.8	212	2 J02227	e antigen precursor
45	964.5	84.0	212	2 S67504	e antigen precursor

ALIGNMENTS

RESULT 1

NKVLAH e antigen precursor / core antigen - hepatitis B virus (subtype ayw4, isolate hb321 and

N/Alternate names: HBe antigen precursor / HBeC antigen; pre-C/C antigen
N/Contains: core antigen; e antigen
C/Species: hepatitis B virus, HBV
A/Variety: subtype ayw4, isolate hb321; isolate patient Ferracuti/83; isolate patient Ca

Chen/ 83
Date: 18-Dec-1981 #sequence, revision 08-Nov-1996 #text change 16-Jul-1999
C/Accession: S47405; S53191; S53209; S53234; S53264; S53249; S53262; S53277; A03711
R/Pluctemiczak, A.
Submitted to the EMBL Data Library, August 1994

A/Description: Molecular cloning and sequencing of two complete genomes of poliovirus isolat
A/Reference number: S47404
A/Accession: S47405
A/Molecule type: DNA
A/Residues: 1-212 <PLU>

A/Cross-references: EMBL:235716; NID:g527435; PIDN:CAA84786.1; PID:g527437
A/Experimental source: subtype ayw4, isolate hb321
R/Lai, M.R.; Mazzoleni, A.P.; Portu, A.; Balserrieri, A.
Submitted to the EMBL Data Library, March 1995

A/Reference number: S53112
A/Accession: S53191
A/Molecule type: DNA
A/Residues: 1-212 <LA1>

A/Cross-references: EMBL:X85283; NID:g736088; PIDN:CAA59593.1; PID:g736090
A/Experimental source: isolate patient Ferracuti/83
A/Accession: S53209
A/Molecule type: DNA
A/Residues: 1-212 <LA2>

A/Cross-references: EMBL:X85290; NID:g736114; PIDN:CAA59609.1; PID:g736116
A/Experimental source: isolate patient Caetlag/83
A/Accession: S53234
A/Molecule type: DNA
A/Residues: 1-212 <LA3>

A/Cross-references: EMBL:X85300; NID:g736150; PIDN:CAA59631.1; PID:g736152
A/Experimental source: isolate patient Sanna/84
A/Accession: S53264
A/Molecule type: DNA
A/Residues: 1-212 <LA4>

A/Cross-references: EMBL:X85313; NID:g736194; PIDN:CAA59659.1; PID:g736196
A/Experimental source: isolate patient Licheri-1/85
A/Accession: S53249
A/Molecule type: DNA
A/Residues: 1-212 <LA5>

A/Cross-references: EMBL:X85306; NID:g736172; PIDN:CAA59644.1; PID:g736174
A/Experimental source: isolate patient Flore-1/86
A/Accession: S53262
A/Molecule type: DNA
A/Residues: 1-212 <LA6>

A/Cross-references: EMBL:X85312; NID:g736191; PIDN:CAA59657.1; PID:g736193
A/Experimental source: isolate patient Licheri/83

A:Accession: S53277
A:Molecule type: DNA
A:Residues: 30-212 <LA7>
A:Cross-references: EMBL:X85317; NID:9736211; PIDN:CAA59669.1; PID:9736214
A:Experimental source: patient Giordo-2'86
A:Note: due to a stop codon between the alternative initiators the e antigen precursor
R:Galibert, F.; Mandart, E.; Filousi, F.; Tiolette, P.; Charnay, P.
A:Title: Nucleotide sequence of the hepatitis B virus genome (subtype ayw) in E. coli.
A:Reference number: A93214; MUID:8101091; PMID:399327
A:Accession: A03711
A:Molecule type: DNA
A:Residues: 1-212 <GAL>
A:Cross-references: GB:J02203; NID:9329640; PIDN:AAA45489.1; PID:9329642
A:Experimental source: subtype ayw
C:Genetics:
A:Gene: C
C:Superfamily: hepatitis B virus core antigen
C:Keywords: alternative initiators; core protein
F:1-29/Domain: signal sequence #status predicted <SIG>
F:30-212/Product: core antigen #status predicted <CAG>
F:179-212/Domain: carboxyl-terminal propeptide #link EAG #status predicted <ECP>
F:179-212/Domain: carboxyl-terminal propeptide #link EAG #status predicted <ECP>
Query Match 90.6%; Score 1040.5; DB 1; Length 212;
Best Local Similarity 97.0%; Pred. No. 2e-82;
Matches 197; Conservative 0; Mismatches 1; Indels 5; Gaps 1;
QY 13 PTVQASKICLGLMGMNDIDPYKEFGATVELLSFLPSDFPSVVDLDTASALYREALESP 72
DB 15 PTVQASKICLGLMGMNDIDPYKEFGATVELLSFLPSDFPSVVDLDTASALYREALESP 74
QY 73 EHCSPHHTALROAILCWEGLMTLATWGVNLEDPDEFGRDASDLVSYNTNMGLKFRQL 132
DB 75 EHCSPHHTALROAILCWEGLMTLATWGVNLEDPDEFGRDASDLVSYNTNMGLKFRQL 129
QY 133 LWFHISCLTFGRRTVEYLVSGVWIRTPPAYRPNAPILSTLPTTVVRRGRSPRRRT 192
DB 130 LWFHISCLTFGRRTVEYLVSGVWIRTPPAYRPNAPILSTLPTTVVRRGRSPRRRT 189
QY 193 PSPRRRSQSPRRRSQSPRRRT 215
DB 190 PSPRRRSQSPRRRSQSPRRRT 212
RESULT 2
S53204
e antigen precursor / core antigen - hepatitis B virus
N:Alternate names: HBe antigen precursor / HBe antigen; pre-C/C antigen
N:Contains: core antigen; e antigen
C:Species: hepatitis B virus, HBV
C:Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 26-Aug-1999
C:Accession: S53204; S53283
R:Lat, M.B.; Mazoleni, A.P.; Porru, A.; Balestrieri, A.
Submitted to the EMBL Data Library, March 1995
A:Reference number: S53112
A:Accession: S53204
A:Molecule type: DNA
A:Residues: 1-212 <LA1>
A:Cross-references: EMBL:X85288; NID:9736107; PIDN:CAA59604.1; PID:9736109
A:Experimental source: isolate patient Casae-1'85
A:Accession: S53283
A:Molecule type: DNA
A:Residues: 1-14, 'A', 16-57, 'D', 59-61, 'N', 63-128, 'R', 130-209, 'S', 211-212 <LA2>
A:Cross-references: EMBL:X85320; NID:9736221; PIDN:CAA59675.1; PID:9736223
A:Experimental source: isolate patient Galibertu 85
C:Genetics:
A:Gene: C
C:Superfamily: hepatitis B virus core antigen
C:Keywords: alternative initiators; core protein
F:1-29/Domain: signal sequence #status predicted <SIG>
F:30-212/Product: core antigen #status predicted <CAG>
F:30-179/Product: e antigen #status predicted <EAG>
F:30-179/Product: e antigen #status predicted <EAG>

F:179-212/Domain: carboxyl-terminal propeptide #link EAG #status predicted <ECP>
Query Match 90.5%; Score 1038.5; DB 2; Length 212;
Best Local Similarity 96.6%; Pred. No. 3e-82;
Matches 196; Conservative 1; Mismatches 1; Indels 5; Gaps 1;
QY 13 PTVQASKICLGLMGMNDIDPYKEFGATVELLSFLPSDFPSVVDLDTASALYREALESP 72
DB 15 PTVQASKICLGLMGMNDIDPYKEFGATVELLSFLPSDFPSVVDLDTASALYREALESP 74
QY 73 EHCSPHHTALROAILCWEGLMTLATWGVNLEDPDEFGRDASDLVSYNTNMGLKFRQL 132
DB 75 EHCSPHHTALROAILCWEGLMTLATWGVNLEDPDEFGRDASDLVSYNTNMGLKFRQL 129
QY 133 LWFHISCLTFGRRTVEYLVSGVWIRTPPAYRPNAPILSTLPTTVVRRGRSPRRRT 192
DB 130 LWFHISCLTFGRRTVEYLVSGVWIRTPPAYRPNAPILSTLPTTVVRRGRSPRRRT 189
QY 193 PSPRRRSQSPRRRSQSPRRRT 215
DB 190 PSPRRRSQSPRRRSQSPRRRT 212
RESULT 3
S53211
e antigen precursor / core antigen - hepatitis B virus (isolate patient Casag-1'85 and
N:Alternate names: HBe antigen; HBe antigen precursor / HBe antigen; pre-C/C antigen
N:Contains: core antigen; e antigen
C:Species: hepatitis B virus, HBV
A:Variety: isolate patient Casag-1'85; isolate patient Ferracuti-1'89
C:Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 26-Aug-1999
C:Accession: S53211; S53197
R:Lat, M.B.; Mazoleni, A.P.; Porru, A.; Balestrieri, A.
Submitted to the EMBL Data Library, March 1995
A:Reference number: S53112
A:Accession: S53211
A:Molecule type: DNA
A:Residues: 1-212 <LA1>
A:Cross-references: EMBL:X85291; NID:9736117; PIDN:CAA59611.1; PID:9736119
A:Experimental source: isolate patient Casag-1'85
A:Accession: S53197
A:Molecule type: DNA
A:Residues: 30-212 <LAW>
A:Cross-references: EMBL:X85284; NID:9736095; PIDN:CAA59596.1; PID:9736098
A:Experimental source: isolate patient Ferracuti-1'89
A:Note: due to a stop codon between the alternative initiators the e antigen precursor
C:Genetics:
A:Gene: C
C:Superfamily: hepatitis B virus core antigen
C:Keywords: alternative initiators; core protein
F:1-29/Domain: signal sequence #status predicted <SIG>
F:30-212/Product: core antigen #status predicted <CAG>
F:30-178/Product: e antigen #status predicted <EAG>
F:179-212/Domain: carboxyl-terminal propeptide #link EAG #status predicted <ECP>
Query Match 90.4%; Score 1037.5; DB 2; Length 212;
Best Local Similarity 96.6%; Pred. No. 3.6e-82;
Matches 196; Conservative 1; Mismatches 1; Indels 5; Gaps 1;
QY 13 PTVQASKICLGLMGMNDIDPYKEFGATVELLSFLPSDFPSVVDLDTASALYREALESP 72
DB 15 PTVQASKICLGLMGMNDIDPYKEFGATVELLSFLPSDFPSVVDLDTASALYREALESP 74
QY 73 EHCSPHHTALROAILCWEGLMTLATWGVNLEDPDEFGRDASDLVSYNTNMGLKFRQL 132
DB 75 EHCSPHHTALROAILCWEGLMTLATWGVNLEDPDEFGRDASDLVSYNTNMGLKFRQL 129
QY 133 LWFHISCLTFGRRTVEYLVSGVWIRTPPAYRPNAPILSTLPTTVVRRGRSPRRRT 192
DB 130 LWFHISCLTFGRRTVEYLVSGVWIRTPPAYRPNAPILSTLPTTVVRRGRSPRRRT 189
QY 193 PSPRRRSQSPRRRSQSPRRRT 215
DB 190 PSPRRRSQSPRRRSQSPRRRT 212

Db 190 PSPRRRSQSPRRRSQSRSC 212

RESULT 4
S32204

e antigen precursor / core antigen - hepatitis B virus (subtype ayw, patients C1000 and N1)
N/Alternate names: HBe antigen precursor / HBe antigen; pre-C/C antigen
N/Contains: core antigen; e antigen
C/Species: hepatitis B virus, HBV
C/Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 20-Oct-2000
C/Accession: S32204; S53207
R/Prepared: Adams, S.; Schleyer, M.J.; Peters, T.; Hettler, F.; Gerok, W.; Rasmack, J.
A/Description: Identification and sequence analysis of hepatitis B virus DNA in immunol
A/Reference number: S32202
A/Accession: S32204
A/Molecule type: DNA

A/Residues: 1-212 <LAT>
A/Cross-references: EMBL:X85289; NID:G736110; PIDN:CAA51257.1; PID:G288930
A/Experimental source: subtype ayw, patient C1000
R/Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
Submitted to the EMBL Data Library, March 1995

A/Reference number: S53112
A/Accession: S53207
A/Molecule type: DNA

A/Residues: 30-212 <LAT>
A/Cross-references: EMBL:X85289; NID:G736110; PIDN:CAA59607.1; PID:G736113

A/Experimental source: isolate patient Castra-2/87
A/Note: due to a stop codon mutation between the alternative initiators the e antigen p

A/Genes: C

C/Superfamily: hepatitis B virus core antigen
C/Keywords: alternative initiators; core protein
F/1-29/Domain: signal sequence #status predicted <SIG>
F/30-212/Product: core antigen #status predicted <CAG>
F/30-178/Product: e antigen #status predicted <EAG>
F/179-212/Domain: carboxyl-terminal propeptide #link EAG #status predicted <ECP>

Query Match 90.4%; Score 1037.5; DB 2; Length 212;
Best Local Similarity 96.6%; Pred. No. 3.6e-82;
Matches 196; Conservative 1; Mismatches 1; Indels 5; Gaps 1;

Qy 13 PTVQASKLCIGWLMGMDIDPYKFGATVLLSFLPSDFPSPVRLDLDTSALYREALSP 72

Db 15 PTVQASKLCIGWLMGMDIDPYKFGATVLLSFLPSDFPSPVRLDLDTSALYREALSP 74

Qy 73 EHCSPHHTALRQAILCWGELMTLATVGVNLDPERFGDASRDLYSVYNTNMGKLFROL 132

Db 75 EHCSPHHTALRQAILCWGELMTLATVGVNLDPERFGDASRDLYSVYNTNMGKLFROL 129

Qy 133 LWFHISCLTFGRETIVLYVSGWIRTPPAYRPNAPILSTLPTTVVRRGRSPRRRT 192

Db 130 LWFHISCLTFGRETIVLYVSGWIRTPPAYRPNAPILSTLPTTVVRRGRSPRRRT 189

Qy 193 PSPRRRSQSPRRRSQSRSC 215

Db 190 PSPRRRSQSPRRRSQSRSC 212

RESULT 5
S20750

e antigen precursor / core antigen - hepatitis B virus (subtype ayw, patient C1)
N/Alternate names: HBe antigen precursor / HBe antigen; pre-C/C antigen
N/Contains: core antigen; e antigen
C/Species: hepatitis B virus, HBV
A/Variety: subtype ayw, patient C1
C/Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C/Accession: S20750

R/Lai, M.E.; Mazzoleni, A.P.; Balestrieri, A.; Meis, A.; Porru, A.
Submitted to the EMBL Data Library, March 1992
A/Description: Sequence analysis of HBV genomes isolated from patients with HBsAg negati
A/Reference number: S20745
A/Accession: S20750

A/Molecule type: DNA

A/Residues: 1-212 <LAT>

A/Cross-references: EMBL:X65258; NID:G59434; PIDN:CAA46354.1; PID:G59436

A/Experimental source: subtype ayw, patient C1

A/Genes: C

C/Superfamily: hepatitis B virus core antigen

C/Keywords: alternative initiators; core protein

F/1-29/Domain: signal sequence #status predicted <SIG>

F/30-212/Product: core antigen #status predicted <CAG>

F/30-178/Product: e antigen #status predicted <EAG>

F/179-212/Domain: carboxyl-terminal propeptide #link EAG #status predicted <ECP>

Query Match 90.4%; Score 1037.5; DB 2; Length 212;
Best Local Similarity 96.6%; Pred. No. 3.6e-82;
Matches 196; Conservative 1; Mismatches 1; Indels 5; Gaps 1;

Qy 13 PTVQASKLCIGWLMGMDIDPYKFGATVLLSFLPSDFPSPVRLDLDTSALYREALSP 72

Db 15 PTVQASKLCIGWLMGMDIDPYKFGATVLLSFLPSDFPSPVRLDLDTSALYREALSP 74

Qy 73 EHCSPHHTALRQAILCWGELMTLATVGVNLDPERFGDASRDLYSVYNTNMGKLFROL 132

Db 75 EHCSPHHTALRQAILCWGELMTLATVGVNLDPERFGDASRDLYSVYNTNMGKLFROL 129

Qy 133 LWFHISCLTFGRETIVLYVSGWIRTPPAYRPNAPILSTLPTTVVRRGRSPRRRT 192

Db 130 LWFHISCLTFGRETIVLYVSGWIRTPPAYRPNAPILSTLPTTVVRRGRSPRRRT 189

Qy 193 PSPRRRSQSPRRRSQSRSC 215

Db 190 PSPRRRSQSPRRRSQSRSC 212

RESULT 6
S53216

e antigen precursor / core antigen - hepatitis B virus (isolate patient Castrag'3)
N/Alternate names: HBe antigen precursor / HBe antigen; pre-C/C antigen
N/Contains: core antigen; e antigen
C/Species: hepatitis B virus, HBV
A/Variety: isolate patient Castrag'3
C/Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 26-Aug-1999
C/Accession: S53216

R/Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
Submitted to the EMBL Data Library, March 1995
A/Reference number: S53112
A/Accession: S53216
A/Molecule type: DNA

A/Residues: 1-212 <LAT>
A/Cross-references: EMBL:X85293; NID:G736124; PIDN:CAA59616.1; PID:G736126
A/Experimental source: isolate patient Castrag'3

A/Genes: C

C/Superfamily: hepatitis B virus core antigen

C/Keywords: alternative initiators; core protein

F/1-29/Domain: signal sequence #status predicted <SIG>

F/30-212/Product: core antigen #status predicted <CAG>

F/30-178/Product: e antigen #status predicted <EAG>

F/179-212/Domain: carboxyl-terminal propeptide #link EAG #status predicted <ECP>

Query Match 90.2%; Score 1035.5; DB 2; Length 212;
Best Local Similarity 96.6%; Pred. No. 5.4e-82;
Matches 196; Conservative 0; Mismatches 2; Indels 5; Gaps 1;

Qy 13 PTVQASKLCIGWLMGMDIDPYKFGATVLLSFLPSDFPSPVRLDLDTSALYREALSP 72

Db 15 PTVQASKLCIGWLMGMDIDPYKFGATVLLSFLPSDFPSPVRLDLDTSALYREALSP 74

Qy 73 EHCSPHHTALRQAILCWGELMTLATVGVNLDPERFGDASRDLYSVYNTNMGKLFROL 132

Db 75 EHCSPHHTALRQAILCWGELMTLATVGVNLDPERFGDASRDLYSVYNTNMGKLFROL 129

Qy 133 LWFHISCLTFGRETIVLYVSGWIRTPPAYRPNAPILSTLPTTVVRRGRSPRRRT 192

Db 130 LMFHISCLTFRGRTVIEYVSGFWIRTPPAYRPPNAPILSTLPTTVRRGRSPRRRT 189
 Qy 193 PPSRRRSQSPRRRSQSPRRSQC 215
 Db 190 PPSRRRSQSPRRRSQSPRRSQC 212

RESULT 7

S53272 e antigen precursor / core antigen - hepatitis B virus (isolate patient Licheri-3'90)

N/Alternate names: HBe antigen precursor / HBeC antigen; pre-C/C antigen
 N/Contains: core antigen; e antigen
 C/Species: hepatitis B virus, HBV

A/Variety: isolate patient Licheri-3'90
 C/Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 26-Aug-1999

C/Accession: S53272
 R/Lat, M.E.: Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
 Submitted to the EMBL Data Library, March 1995

A/Reference number: S53112
 A/Accession: S53272

A/Molecule type: DNA

A/Residues: 1-212 <LAT>

A/Cross-references: EMBL:X85315; NID:G736205; PIDN:CAAS9664.1; PID:G736207

A/Experimental source: isolate patient Licheri-3'90

C/Genetics:

A/Gene: C

C/Superfamily: hepatitis B virus core antigen

C/Keywords: alternative initiators; core protein

F/1-29/Domain: signal sequence #status predicted <SIG>

F/30-212/Product: core antigen #status predicted <CAG>

F/30-178/Product: e antigen #status predicted <EAG>

F/179-212/Domain: carboxyl-terminal propeptide #link EAG #status predicted <ECP>

Query Match 90.1%; Score 1035.5; DB 2; Length 212;

Best Local Similarity 96.6%; Pred. No. 5.4e-82; Indels 5; Gaps 1;

Matches 196; Conservative 0; Mismatches 2; Indels 5; Gaps 1;

Qy 13 PTVOASKLCIGWLMGMDIDPYKEFGATVELLSFLPSDFPSPVVDLDTASALYREALESP 72
 Db 15 PTVOASKLCIGWLMGMDIDPYKEFGATVELLSFLPSDFPSPVVDLDTASALYREALESP 74

Qy 73 EHCSPHHTALROAILCWGELMTLATWGVNLEDPFRGASRDLYSVYNTNGLKFRQL 132
 Db 75 EHCSPHHTALROAILCWGELMTLATWGVNLEDPFRGASRDLYSVYNTNGLKFRQL 129

Qy 133 LMFHISCLTFRGRTVIEYVSGFWIRTPPAYRPPNAPILSTLPTTVRRGRSPRRRT 192
 Db 130 LMFHISCLTFRGRTVIEYVSGFWIRTPPAYRPPNAPILSTLPTTVRRGRSPRRRT 189

Qy 193 PPSRRRSQSPRRRSQSPRRSQC 215
 Db 190 PPSRRRSQSPRRRSQSPRRSQC 212

Qy 193 PPSRRRSQSPRRRSQSPRRSQC 215
 Db 190 PPSRRRSQSPRRRSQSPRRSQC 212

Qy 193 PPSRRRSQSPRRRSQSPRRSQC 215
 Db 190 PPSRRRSQSPRRRSQSPRRSQC 212

Qy 193 PPSRRRSQSPRRRSQSPRRSQC 215
 Db 190 PPSRRRSQSPRRRSQSPRRSQC 212

Qy 193 PPSRRRSQSPRRRSQSPRRSQC 215
 Db 190 PPSRRRSQSPRRRSQSPRRSQC 212

Qy 193 PPSRRRSQSPRRRSQSPRRSQC 215
 Db 190 PPSRRRSQSPRRRSQSPRRSQC 212

Qy 193 PPSRRRSQSPRRRSQSPRRSQC 215
 Db 190 PPSRRRSQSPRRRSQSPRRSQC 212

Qy 193 PPSRRRSQSPRRRSQSPRRSQC 215
 Db 190 PPSRRRSQSPRRRSQSPRRSQC 212

Qy 193 PPSRRRSQSPRRRSQSPRRSQC 215
 Db 190 PPSRRRSQSPRRRSQSPRRSQC 212

Qy 193 PPSRRRSQSPRRRSQSPRRSQC 215
 Db 190 PPSRRRSQSPRRRSQSPRRSQC 212

Qy 193 PPSRRRSQSPRRRSQSPRRSQC 215
 Db 190 PPSRRRSQSPRRRSQSPRRSQC 212

Qy 193 PPSRRRSQSPRRRSQSPRRSQC 215
 Db 190 PPSRRRSQSPRRRSQSPRRSQC 212

Qy 193 PPSRRRSQSPRRRSQSPRRSQC 215
 Db 190 PPSRRRSQSPRRRSQSPRRSQC 212

A/Gene: C
 C/Superfamily: hepatitis B virus core antigen
 C/Keywords: alternative initiators; core protein
 F/1-29/Domain: signal sequence #status predicted <SIG>
 F/30-212/Product: core antigen #status predicted <CAG>
 F/30-178/Product: e antigen #status predicted <EAG>
 F/179-212/Domain: carboxyl-terminal propeptide #link EAG #status predicted <ECP>

Query Match

Best Local Similarity 96.1%; Score 1034.5; DB 2; Length 212;

Matches 195; Conservative 2; Mismatches 1; Indels 5; Gaps 1;

Qy 13 PTVOASKLCIGWLMGMDIDPYKEFGATVELLSFLPSDFPSPVVDLDTASALYREALESP 72
 Db 15 PTVOASKLCIGWLMGMDIDPYKEFGATVELLSFLPSDFPSPVVDLDTASALYREALESP 74

Qy 73 EHCSPHHTALROAILCWGELMTLATWGVNLEDPFRGASRDLYSVYNTNGLKFRQL 132
 Db 75 EHCSPHHTALROAILCWGELMTLATWGVNLEDPFRGASRDLYSVYNTNGLKFRQL 129

Qy 133 LMFHISCLTFRGRTVIEYVSGFWIRTPPAYRPPNAPILSTLPTTVRRGRSPRRRT 192
 Db 130 LMFHISCLTFRGRTVIEYVSGFWIRTPPAYRPPNAPILSTLPTTVRRGRSPRRRT 189

Qy 193 PPSRRRSQSPRRRSQSPRRSQC 215
 Db 190 PPSRRRSQSPRRRSQSPRRSQC 212

Qy 193 PPSRRRSQSPRRRSQSPRRSQC 215
 Db 190 PPSRRRSQSPRRRSQSPRRSQC 212

Qy 193 PPSRRRSQSPRRRSQSPRRSQC 215
 Db 190 PPSRRRSQSPRRRSQSPRRSQC 212

Qy 193 PPSRRRSQSPRRRSQSPRRSQC 215
 Db 190 PPSRRRSQSPRRRSQSPRRSQC 212

Qy 193 PPSRRRSQSPRRRSQSPRRSQC 215
 Db 190 PPSRRRSQSPRRRSQSPRRSQC 212

Qy 193 PPSRRRSQSPRRRSQSPRRSQC 215
 Db 190 PPSRRRSQSPRRRSQSPRRSQC 212

Qy 193 PPSRRRSQSPRRRSQSPRRSQC 215
 Db 190 PPSRRRSQSPRRRSQSPRRSQC 212

Qy 193 PPSRRRSQSPRRRSQSPRRSQC 215
 Db 190 PPSRRRSQSPRRRSQSPRRSQC 212

Qy 193 PPSRRRSQSPRRRSQSPRRSQC 215
 Db 190 PPSRRRSQSPRRRSQSPRRSQC 212

Qy 193 PPSRRRSQSPRRRSQSPRRSQC 215
 Db 190 PPSRRRSQSPRRRSQSPRRSQC 212

Qy 193 PPSRRRSQSPRRRSQSPRRSQC 215
 Db 190 PPSRRRSQSPRRRSQSPRRSQC 212

Qy 193 PPSRRRSQSPRRRSQSPRRSQC 215
 Db 190 PPSRRRSQSPRRRSQSPRRSQC 212

Qy 193 PPSRRRSQSPRRRSQSPRRSQC 215
 Db 190 PPSRRRSQSPRRRSQSPRRSQC 212

Qy 193 PPSRRRSQSPRRRSQSPRRSQC 215
 Db 190 PPSRRRSQSPRRRSQSPRRSQC 212

Qy 193 PPSRRRSQSPRRRSQSPRRSQC 215
 Db 190 PPSRRRSQSPRRRSQSPRRSQC 212

Qy 193 PPSRRRSQSPRRRSQSPRRSQC 215
 Db 190 PPSRRRSQSPRRRSQSPRRSQC 212

Qy 193 PPSRRRSQSPRRRSQSPRRSQC 215
 Db 190 PPSRRRSQSPRRRSQSPRRSQC 212

Qy 193 PPSRRRSQSPRRRSQSPRRSQC 215
 Db 190 PPSRRRSQSPRRRSQSPRRSQC 212

Qy 193 PPSRRRSQSPRRRSQSPRRSQC 215
 Db 190 PPSRRRSQSPRRRSQSPRRSQC 212

Qy 193 PPSRRRSQSPRRRSQSPRRSQC 215
 Db 190 PPSRRRSQSPRRRSQSPRRSQC 212

Qy 193 PPSRRRSQSPRRRSQSPRRSQC 215
 Db 190 PPSRRRSQSPRRRSQSPRRSQC 212

Qy 193 PPSRRRSQSPRRRSQSPRRSQC 215
 Db 190 PPSRRRSQSPRRRSQSPRRSQC 212

Qy 193 PPSRRRSQSPRRRSQSPRRSQC 215
 Db 190 PPSRRRSQSPRRRSQSPRRSQC 212

Db 190 PPSRRRSQSPRRRSQSRSSQC 212

RESULT 10

S53163

e antigen precursor / core antigen - hepatitis B virus (isolate patient Vittorina '92)
N:Alternate names: HBe antigen precursor / HBeC antigen; pre-C/C antigen
N:Contains: core antigen; e antigen
C:Species: hepatitis B virus, HBV

A:Variety: isolate patient Vittorina '92
C:Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 26-Aug-1999
C:Accession: S53163

R:Lat, M.R.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
Submitted to the EMBL Data Library, March 1995

A:Reference number: S53112

A:Accession: S53163

A:Molecule type: DNA

A:Residues: 1-212 <LAI>

A:Cross-references: EMBL:X65256; NID:G736050; PIDN:CAA59519.1; PID:G736052

A:Experimental source: isolate patient Vittorina '92

C:Genetics:

A:Gene: C

C:Superfamily: hepatitis B virus core antigen

C:Keywords: alternative initiators; core protein

F:1-29/Domain: signal sequence #status predicted <SIG>

F:30-212/Product: core antigen #status predicted <CAG>

F:179-212/Domain: carboxyl-terminal propeptide #link EAG #status predicted <BCP>

Query Match 90.1%; Score 1034.5; DB 2; Length 212;
Best Local Similarity 96.6%; Pred. No. 6 6e-82; Indels 5; Gaps 1;

Matches 196; Conservative 0; Mismatches 2; Indels 5; Gaps 1;

13 PTVQASKLCGLMGMDIDPYKFGATVELLSFLPSDFPSVRDLDTASALYREALESP 72

15 PTVQASKLCGLMGMDIDPYKFGATVELLSFLPSDFPSVRDLDTASALYREALESP 74

73 EHCSPHHTALROAILCWGELMTLATVGVNLEDPERFSDASDLYSVYNTMGLKFRQL 132

75 EHCSPHHTALROAILCWGELMTLATVGVNLEDPERFSDASDLYSVYNTMGLKFRQL 129

133 LMFHISCLTFGRETIVLYVSGWIRTPPAYRPNAPILSTLPTTVVRRGRSPRRRT 192

130 LMFHISCLTFGRETIVLYVSGWIRTPPAYRPNAPILSTLPTTVVRRGRSPRRRT 189

193 PPSRRRSQSPRRRSQSRSSQC 215

190 PPSRRRSQSPRRRSQSRSSQC 212

Db 190 PPSRRRSQSPRRRSQSRSSQC 212

RESULT 11

S53251

e antigen precursor / core antigen - hepatitis B virus (isolate patient Fiore '85)
N:Alternate names: HBe antigen precursor / HBeC antigen; pre-C/C antigen
N:Contains: core antigen; e antigen
C:Species: hepatitis B virus, HBV

A:Variety: isolate patient Fiore '85

C:Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 26-Aug-1999

C:Accession: S53251

R:Lat, M.R.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
Submitted to the EMBL Data Library, March 1995

A:Reference number: S53112

A:Accession: S53251

A:Molecule type: DNA

A:Residues: 1-212 <LAI>

A:Cross-references: EMBL:X65307; NID:G736175; PIDN:CAA59646.1; PID:G736177

A:Experimental source: isolate patient Fiore '85

C:Genetics:

A:Gene: C

C:Superfamily: hepatitis B virus core antigen

C:Keywords: alternative initiators; core protein

F:1-29/Domain: signal sequence #status predicted <SIG>

F:30-212/Product: core antigen #status predicted <CAG>

F:30-178/Product: e antigen #status predicted <ENG>
F:179-212/Domain: carboxyl-terminal propeptide #link EAG #status predicted <BCP>

Query Match 89.9%; Score 1031.5; DB 2; Length 212;
Best Local Similarity 96.1%; Pred. No. 1.2e-81; Indels 5; Gaps 1;

Matches 195; Conservative 1; Mismatches 2; Indels 5; Gaps 1;

13 PTVQASKLCGLMGMDIDPYKFGATVELLSFLPSDFPSVRDLDTASALYREALESP 72

15 PTVQASKLCGLMGMDIDPYKFGATVELLSFLPSDFPSVRDLDTASALYREALESP 74

73 EHCSPHHTALROAILCWGELMTLATVGVNLEDPERFSDASDLYSVYNTMGLKFRQL 132

75 EHCSPHHTALROAILCWGELMTLATVGVNLEDPERFSDASDLYSVYNTMGLKFRQL 129

133 LMFHISCLTFGRETIVLYVSGWIRTPPAYRPNAPILSTLPTTVVRRGRSPRRRT 192

130 LMFHISCLTFGRETIVLYVSGWIRTPPAYRPNAPILSTLPTTVVRRGRSPRRRT 189

193 PPSRRRSQSPRRRSQSRSSQC 215

190 PPSRRRSQSPRRRSQSRSSQC 212

Db 190 PPSRRRSQSPRRRSQSRSSQC 212

RESULT 12

S20746

e antigen precursor / core antigen - hepatitis B virus (subtype ayw, patient C)
N:Alternate names: HBe antigen precursor / HBeC antigen; pre-C/C antigen
N:Contains: core antigen; e antigen
C:Species: hepatitis B virus, HBV

A:Variety: subtype ayw, patient C

C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999

C:Accession: S20746

R:Lat, M.R.; Mazzoleni, A.P.; Balestrieri, A.; Mellis, A.; Porru, A.
Submitted to the EMBL Data Library, March 1992

A:Description: Sequence analysis of HBV genomes isolated from patients with HBsAg negativ

A:Reference number: S20745

A:Accession: S20745

A:Molecule type: DNA

A:Residues: 1-212 <LAI>

A:Cross-references: EMBL:X65257; NID:G59429; PIDN:CAA46350.1; PID:G59431

A:Experimental source: subtype ayw, patient C

C:Genetics:

A:Gene: C

C:Superfamily: hepatitis B virus core antigen

C:Keywords: alternative initiators; core protein

F:1-29/Domain: signal sequence #status predicted <SIG>

F:30-212/Product: core antigen #status predicted <CAG>

F:179-212/Domain: carboxyl-terminal propeptide #link EAG #status predicted <BCP>

Query Match 89.9%; Score 1031.5; DB 2; Length 212;
Best Local Similarity 96.1%; Pred. No. 1.2e-81; Indels 5; Gaps 1;

Matches 195; Conservative 1; Mismatches 2; Indels 5; Gaps 1;

13 PTVQASKLCGLMGMDIDPYKFGATVELLSFLPSDFPSVRDLDTASALYREALESP 72

15 PTVQASKLCGLMGMDIDPYKFGATVELLSFLPSDFPSVRDLDTASALYREALESP 74

73 EHCSPHHTALROAILCWGELMTLATVGVNLEDPERFSDASDLYSVYNTMGLKFRQL 132

75 EHCSPHHTALROAILCWGELMTLATVGVNLEDPERFSDASDLYSVYNTMGLKFRQL 129

133 LMFHISCLTFGRETIVLYVSGWIRTPPAYRPNAPILSTLPTTVVRRGRSPRRRT 192

130 LMFHISCLTFGRETIVLYVSGWIRTPPAYRPNAPILSTLPTTVVRRGRSPRRRT 189

193 PPSRRRSQSPRRRSQSRSSQC 215

190 PPSRRRSQSPRRRSQSRSSQC 212

Db 190 PPSRRRSQSPRRRSQSRSSQC 212

RESULT 13

```

553159
e antigen precursor / core antigen - hepatitis B virus (isolate patient Garrucci'u'90)
N/Alternate names: HBe antigen precursor / HBe antigen; pre-C/C antigen
N/Contains: core antigen; e antigen
C/Species: hepatitis B virus, HBV
A/Variety: isolate patient Garrucci'u'90
C/Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 26-Aug-1999
A/Accession: S53159
R/Lat: M.E.; Mazzoleni, A.P.; Portu, A.; Balestrieri, A.
submitted to the EMBL Data Library, March 1995
A/Reference number: S53112
A/Accession: S53159
A/Molecule type: DNA
A/Residues: 1-212 <Lat>
A/Cross-references: EMBL:X85286; NID:G736044; PIDN:CAAS9563.1; PID:G736046
A/Experimental source: isolate patient Garrucci'u'90
C/Genetics:
A/Gene: C
C/Superfamily: hepatitis B virus core antigen
C/Keywords: alternative initiators; core protein
F/1-29/Domain: signal sequence #status predicted <Sig>
F/30-212/Product: core antigen #status predicted <CAG>
F/30-178/Product: e antigen #status predicted <EAG>
F/179-212/Domain: carboxyl-terminal propeptide #link EAG #status predicted <ECP>

Query Match      89.8%; Score 1030.5; DB 2; Length 212;
Best Local Similarity 95.6%; Pred. No. 1,5e-81;
Matches 194; Conservative 3; Mismatches 1; Indels 5; Gaps 1;

Qy 13 PTVOASRLCLGMIWGMIDIDYKEFGATVELLSFLSPDFPSVDLDTASALYREALESP 72
Db 15 PTVOASRLCLGMIWGMIDIDYKEFGATVELLSFLSPDFPSVDLDTASALYREALESP 74

Qy 73 EHCSPHHTALRQAILKCEGLMTLATWGVNLEDERFGDASRDLYSVYNTNMGLKFEROL 132
Db 75 EHCSPHHTALRQAILKCEGLMTLATWGVNLEDERFGDASRDLYSVYNTNMGLKFEROL 129

Qy 133 LWFHISCLTGREIVIVLVSGVWIRTPPAYRPPNAPILSTLEPTVTVVRRRSGSPRRRT 192
Db 130 LWFHISCLTGREIVIVLVSGVWIRTPPAYRPPNAPILSTLEPTVTVVRRRSGSPRRRT 189

Qy 193 PSPRRRRSOSPFRRRRSQSRPEOC 215
Db 190 PSPRRRRSOSPFRRRRSQSRPEOC 212

RESULT 14
S53200
e antigen precursor / core antigen - hepatitis B virus (isolate patient Ferracuti-3'91)
N/Alternate names: HBe antigen precursor / HBe antigen; pre-C/C antigen
N/Contains: core antigen; e antigen
C/Species: hepatitis B virus, HBV
A/Variety: isolate patient Ferracuti-3'91
C/Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 26-Aug-1999
A/Accession: S53200
R/Lat: M.E.; Mazzoleni, A.P.; Portu, A.; Balestrieri, A.
submitted to the EMBL Data Library, March 1995
A/Reference number: S53112
A/Accession: S53200
A/Molecule type: DNA
A/Residues: 1-212 <Lat>
A/Cross-references: EMBL:X85286; NID:G736101; PIDN:CAAS9600.1; PID:G736103
A/Experimental source: isolate patient Ferracuti-3'91
C/Genetics:
A/Gene: C
C/Superfamily: hepatitis B virus core antigen
C/Keywords: alternative initiators; core protein
F/1-29/Domain: signal sequence #status predicted <Sig>
F/30-212/Product: core antigen #status predicted <CAG>
F/30-178/Product: e antigen #status predicted <EAG>
F/179-212/Domain: carboxyl-terminal propeptide #link EAG #status predicted <ECP>

Query Match      89.8%; Score 1030.5; DB 2; Length 212;

```

	Best Local Similarity	96.1%	Fried. No. 1.5e-81						
	Matches	195	Conservative	1	Mismatches	2	Indels	5	Gaps
QY	13	PTVASKKLCGMLMGMDIDPYKEFGATVELLSFLPSDFPFSSVDRLDTASALYREALESP	72						
Db	15	PTVASKKLCGLMGMDIDPYKEFGATVELLSFLPSDFPFSSVDRLDTASALYREALESP	74						
QY	73	EHCSPHTALFQALLCWEGLMTLTATYGVGNLEDEDFRGSARDLVSVVNNMGKFRQL	132						
Db	75	EHCSPHTALFQALLCWEGLMTLTATYGVGNLEDEDFRGSARDLVSVVNNMGKFRQL	129						
QY	133	LMFHSICLTFERETVEYLVSGVWIRTPAYRPNNAPILSTLPETTYVRGGSPPRRT	192						
Db	130	LMFHSICLTFERETVEYLVSGVWIRTPAYRPNNAPILSTLPETTYVRGGSPPRRT	189						
QY	193	PSPRRRRSQSPRRRRSQSREPOC	215						
Db	190	PSPRRRRSQSPRRRRSQSREPOC	212						

```

RESULT 15
SS3202
e antigen precursor / core antigen - hepatitis B virus (Isolate patient Caesca'83)
N:Alternate names: HBe antigen precursor / HBe antigen; pre-C/C antigen
N:Contains: core antigen; e antigen
C:Species: hepatitis B virus, HBV
C:Varyety: Isolate patient Caesca'83
C:Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 26-Aug-1999
C:Accession: SS3202
R:Lat, M.E.; Mazzoleni, A.P.; Porru, A.; Balistreri, A.
Submitted to the EMBL Data Library, March 1995
A:Reference number: SS3112
A:Accession: SS3202
M:Molecule type: DNA
A:Residues: 1-212 <LAI>
A:Cross-References: EMBL:X85287; NID:G736104; PIDN:CAAS9602.1; PID:G736106
A:Experimental source: Isolate patient Caesca'83
C:Genetics:
A:Gene: C
C:Superfamily: hepatitis B virus core antigen
C:Keywords: alternative initiators / core protein
F:1-20/Domain: signal sequence #status predicted <SIG>
F:30-212/Product: core antigen #status predicted <CAG>
F:30-178/Product: e antigen #status predicted <EAG>
F:179-212/Domain: carboxyl-terminal propeptide #link EAG #status predicted <ECP>

```

	Query Match	89.8%	Score 1030.5;	DB 2;	Length 212;
	Best Local Similarity	96.1%;	Pred No. 1.5e-81;		
	Matches 195;	Conservative	1; Mismatches	2; Indels	5; Gaps
Qy	13	PTVQASKLCLGIMGMGDIDPEYKEFGATVELLSFLPEDFPFSYRDLDTLSALYREALSESP	72		
Db	15	PTYQASLCLGIMGMGDIDPEYKEFGATVELLSFLPEDFPFSVALDLSALYREALSESP	74		
Qy	73	EKCSPHHTALAROAILLCMGEMLMTLATVGVNLDEPPEFGRGASRDLYSVYNTNNGLEFRQL	132		
Db	75	EHCSPHHTAROAILLCEGLMNTLATVGVNLDEP-----ASRLVSYVYNTNNGLEFRQL	129		
Qy	133	LWFHISCLTGREIVIEIYLVSPGVWIRTPPAYAPPNPAPIITSLPETTTVVARRRRRSRRRT	192		
Db	130	LMFHISCLTGREIVIEIYLVSPGVWIRTPPAYAPPNPAPILSTLPETTVVARRRRSRPRKI	189		
Qy	193	PSPPRRRSQSPPRRRSQRSEPOC	215		
Db	190	PSPPRRRSQSPPRRRSQRSRESC	212		

Search completed: February 3, 2003, 09:26:28
Job time : 15.2334 secs

Query Match 89.8%; Score 1030.5; DB 2; Length 212;

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 09:24:27 ; Search time 17.6045 Seconds
(without alignments)
988.252 Million cell updates/sec

Title: US-09-890-752a-2

Perfect score: 1148
Sequence: 1 MFLSIFSRIGDPTVQASKL.....RRRSQSPRRRSQSPRQC 215

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 404799 seqs, 80919614 residues

Total number of hits satisfying chosen parameters: 404799

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents AA New:*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	951.5	82.9	183	US-10-274-616-1	Sequence 1, Appl
2	932.5	81.2	183	US-10-274-616-4	Sequence 4, Appl
3	919.5	80.1	185	US-10-274-616-2	Sequence 2, Appl
4	911.5	79.4	185	US-10-274-616-3	Sequence 3, Appl
5	901.5	78.5	183	US-10-226-007-31	Sequence 31, Appl
6	754.5	65.7	217	US-10-274-616-6	Sequence 6, Appl
7	722.5	62.9	169	PCT-US02-17536-17	Sequence 17, Appl
8	722.5	62.9	169	PCT-US02-17536-18	Sequence 18, Appl
9	661	57.6	183	US-10-274-616-5	Sequence 5, Appl
10	114	9.9	20	US-09-863-054-21	Sequence 21, Appl
11	113.5	9.9	312	US-10-218-140-1758	Sequence 1758, Ap
12	110	9.6	20	US-09-863-054-19	Sequence 19, Appl
13	110	9.6	20	US-10-226-007-714	Sequence 714, App
14	109	9.5	20	US-10-226-007-646	Sequence 646, App
15	109	9.5	20	US-10-226-007-679	Sequence 679, App
16	109	9.5	20	US-10-226-007-688	Sequence 688, App
17	109	9.5	20	US-10-226-007-696	Sequence 696, App
18	109	9.5	20	US-10-226-007-703	Sequence 703, App
19	109	9.5	20	US-10-226-007-709	Sequence 709, App
20	108	9.4	20	US-10-226-007-669	Sequence 669, App
21	108	9.4	20	US-10-226-007-718	Sequence 718, App
22	106	9.2	20	US-10-226-007-658	Sequence 658, App
23	105	9.1	19	US-10-226-007-678	Sequence 678, App
24	105	9.1	19	US-10-226-007-695	Sequence 695, App
25	105	9.1	19	US-10-226-007-687	Sequence 687, App
26	105	9.1	19	US-10-226-007-708	Sequence 708, App

27	105	9.1	20	US-10-226-007-721	Sequence 721, App
28	104	9.1	19	US-10-226-007-713	Sequence 713, App
29	104	9.1	20	US-09-863-054-17	Sequence 17, Appl
30	104	9.1	20	US-10-226-007-607	Sequence 607, App
31	104	9.1	20	US-10-226-007-620	Sequence 620, App
32	103.5	9.0	32	US-10-326-908-12	Sequence 12, Appl
33	103	9.0	19	US-10-226-007-668	Sequence 668, App
34	103	9.0	20	US-10-226-007-594	Sequence 594, App
35	103	9.0	20	US-10-226-007-594	Sequence 594, App
36	103	9.0	20	US-10-226-007-633	Sequence 633, App
37	102.5	8.9	24	US-10-326-908-11	Sequence 723, App
38	102	8.9	19	US-10-226-007-645	Sequence 645, App
39	102	8.9	19	US-10-226-007-657	Sequence 657, App
40	102	8.9	19	US-10-226-007-702	Sequence 702, App
41	101	8.8	18	US-10-226-007-677	Sequence 677, App
42	101	8.8	18	US-10-226-007-685	Sequence 685, App
43	101	8.8	19	US-10-226-007-717	Sequence 717, App
44	100	8.7	20	US-09-865-294A-20	Sequence 20, Appl
45	100	8.7	20	US-10-226-007-724	Sequence 724, App

ALIGNMENTS

RESULT 1
US-10-274-616-1
Sequence 1, Application US/10274616
GENERAL INFORMATION:
APPLICANT: Birkett, Ashley J.
TITLE OF INVENTION: INFLUENZA IMMUNOGEN AND VACCINE
FILE REFERENCE: ICC 127.0 4564/89545
CURRENT APPLICATION NUMBER: US/10/274,616
CURRENT FILING DATE: 2002-10-21
PRIOR APPLICATION NUMBER: 09/930,915
PRIOR FILING DATE: 2001-08-15
PRIOR APPLICATION NUMBER: 10/080,299
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 87
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 183
TYPE: PRT
ORGANISM: Hepatitis B virus
US-10-274-616-1

Query Match 82.9%; Score 951.5; DB 6; Length 183;
Best Local Similarity 96.8%; Pred. No. 7.4e-62;

Matches 182; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

QY 28 MDIDPYKFGATVBLISFLPSDFPVRDLDTASALVREALESPHCSPHHTALQAIL 87
DB 1 MDIDPYKFGATVBLISFLPSDFPVRDLDTASALVREALESPHCSPHHTALQAIL 60
QY 88 CWGELMTLWVGNLEDEFRGDASRDVSVYNTNMGLKFRQLMFIISCTIFGRFV 147
DB 61 CWGELMTLWVGNLEDEFRGDASRDVSVYNTNMGLKFRQLMFIISCTIFGRFV 115
QY 148 IETLVGFWIRPPVPRPNAPILSTPEYTVRRRGSPPRRTPSPRRRSQSPRRR 207
DB 116 IETLVGFWIRPPVPRPNAPILSTPEYTVRRRGSPPRRTPSPRRRSQSPRRR 175
QY 208 SQGRSQC 215
DB 176 SQGRSQC 183

RESULT 2
US-10-274-616-4
Sequence 4, Application US/10274616
GENERAL INFORMATION:
APPLICANT: Birkett, Ashley J.
TITLE OF INVENTION: INFLUENZA IMMUNOGEN AND VACCINE
FILE REFERENCE: ICC 127.0 4564/89545

CURRENT APPLICATION NUMBER: US/10/274,616
 CURRENT FILING DATE: 2002-10-21
 PRIOR APPLICATION NUMBER: 09/930,915
 PRIOR FILING DATE: 2001-08-15
 PRIOR APPLICATION NUMBER: 10/080,299
 PRIOR FILING DATE: 2002-02-21
 NUMBER OF SEQ ID NOS: 87
 SOFTWARE: Patent in version 3.1
 SEQ ID NO 4
 LENGTH: 183
 TYPE: PRT
 ORGANISM: Hepatitis B virus
 US-10-274-616-4

Query Match 81.2%; Score 932.5; DB 6; Length 183;
 Best Local Similarity 93.6%; Pred. No. 1,8e-60;
 Matches 176; Conservative 5; Mismatches 2; Indels 5; Gaps 1;

QY 28 MDIDPYKEGATVELLSFLPSDFPSVVDLDTASALYREALSPHCHSPHHTALRQAIL 87
 DB 1 MDIDPYKEGATVELLSFLPSDFPSVVDLDTASALYREALSPHCHSPHHTALRQAIL 60
 QY 88 CWGELMTLATWGVNLEDPFRGDASRDIVSYNTNMGKXPROLWFMHISCLTFGRGV 147
 DB 61 CWGELMTLATWGVNLEDPFRGDASRDIVSYNTNMGKXPROLWFMHISCLTFGRGV 115
 QY 148 IEYVSGFWIRTPPAYRPPNAPILSTLPETTVRRRGSPPRRTPSPRRRSQSPRR 207
 DB 116 IEYVSGFWIRTPPAYRPPNAPILSTLPETTVRRRGSPPRRTPSPRRRSQSPRR 175
 QY 208 RRSQSPRR 215
 DB 176 RRSQSPRR 183

RESULT 3
 US-10-274-616-2
 Sequence 2, Application US/10274616
 GENERAL INFORMATION:
 APPLICANT: Birkett, Ashley J.
 TITLE OF INVENTION: INFLUENZA IMMUNOGEN AND VACCINE
 FILE REFERENCE: ICC 127.0 4564/88545
 CURRENT APPLICATION NUMBER: US/10/274,616
 CURRENT FILING DATE: 2002-10-21
 PRIOR APPLICATION NUMBER: 09/930,915
 PRIOR FILING DATE: 2001-08-15
 PRIOR APPLICATION NUMBER: 10/080,299
 PRIOR FILING DATE: 2002-02-21
 NUMBER OF SEQ ID NOS: 87
 SOFTWARE: Patent in version 3.1
 SEQ ID NO 2
 LENGTH: 185
 TYPE: PRT
 ORGANISM: Hepatitis B virus
 US-10-274-616-2

Query Match 80.1%; Score 919.5; DB 6; Length 185;
 Best Local Similarity 93.2%; Pred. No. 1,5e-59;
 Matches 177; Conservative 3; Mismatches 3; Indels 7; Gaps 2;

QY 28 MDIDPYKEGATVELLSFLPSDFPSVVDLDTASALYREALSPHCHSPHHTALRQAIL 87
 DB 1 MDIDPYKEGATVELLSFLPSDFPSVVDLDTASALYREALSPHCHSPHHTALRQAIL 60
 QY 88 CWGELMTLATWGVNLEDPFRGDASRDIVSYNTNMGKXPROLWFMHISCLTFGRGV 147
 DB 61 CWGELMTLATWGVNLEDPFRGDASRDIVSYNTNMGKXPROLWFMHISCLTFGRGV 115
 QY 148 IEYVSGFWIRTPPAYRPPNAPILSTLPETTVRRRGSPPRRTPSPRRRSQSPRR 205
 DB 116 IEYVSGFWIRTPPAYRPPNAPILSTLPETTVRRRGSPPRRTPSPRRRSQSPRR 175
 QY 206 RRSQSPRR 215

DB 176 RRSQSPRR 185

RESULT 4
 US-10-274-616-3
 Sequence 3, Application US/10274616
 GENERAL INFORMATION:
 APPLICANT: Birkett, Ashley J.
 TITLE OF INVENTION: INFLUENZA IMMUNOGEN AND VACCINE
 FILE REFERENCE: ICC 127.0 4564/88545
 CURRENT APPLICATION NUMBER: US/10/274,616
 CURRENT FILING DATE: 2002-10-21
 PRIOR APPLICATION NUMBER: 09/930,915
 PRIOR FILING DATE: 2001-08-15
 PRIOR APPLICATION NUMBER: 10/080,299
 PRIOR FILING DATE: 2002-02-21
 NUMBER OF SEQ ID NOS: 87
 SOFTWARE: Patent in version 3.1
 SEQ ID NO 3
 LENGTH: 185
 TYPE: PRT
 ORGANISM: Hepatitis B virus
 US-10-274-616-3

Query Match 79.4%; Score 911.5; DB 6; Length 185;
 Best Local Similarity 92.6%; Pred. No. 5,8e-59;
 Matches 176; Conservative 3; Mismatches 4; Indels 7; Gaps 2;

QY 28 MDIDPYKEGATVELLSFLPSDFPSVVDLDTASALYREALSPHCHSPHHTALRQAIL 87
 DB 1 MDIDPYKEGATVELLSFLPSDFPSVVDLDTASALYREALSPHCHSPHHTALRQAIL 60
 QY 88 CWGELMTLATWGVNLEDPFRGDASRDIVSYNTNMGKXPROLWFMHISCLTFGRGV 147
 DB 61 CWGELMTLATWGVNLEDPFRGDASRDIVSYNTNMGKXPROLWFMHISCLTFGRGV 115
 QY 148 IEYVSGFWIRTPPAYRPPNAPILSTLPETTVRRRGSPPRRTPSPRRRSQSPRR 205
 DB 116 IEYVSGFWIRTPPAYRPPNAPILSTLPETTVRRRGSPPRRTPSPRRRSQSPRR 175
 QY 206 RRSQSPRR 215
 DB 176 RRSQSPRR 185

RESULT 5
 US-10-226-007-31
 Sequence 31, Application US/10226007
 GENERAL INFORMATION:
 APPLICANT: Myriad Genetics, Inc.
 APPLICANT: Morham, Scott
 APPLICANT: Zavit, Kenton
 APPLICANT: Hobden, Adrian
 TITLE OF INVENTION: Compositions and Therapeutic Methods for Viral Infection
 FILE REFERENCE: 5005, 01
 CURRENT APPLICATION NUMBER: US/10/226,007
 CURRENT FILING DATE: 2002-11-15
 PRIOR APPLICATION NUMBER: US 60/313,883
 PRIOR FILING DATE: 2001-08-21
 NUMBER OF SEQ ID NOS: 1673
 SOFTWARE: Patent in version 3.1
 SEQ ID NO 31
 LENGTH: 183
 TYPE: PRT
 ORGANISM: Hepatitis B virus
 US-10-226-007-31

Query Match 78.5%; Score 901.5; DB 6; Length 183;
 Best Local Similarity 90.4%; Pred. No. 3,1e-58;
 Matches 170; Conservative 8; Mismatches 5; Indels 5; Gaps 1;

QY 28 MDIDPYKEGATVELLSFLPSDFPSVVDLDTASALYREALSPHCHSPHHTALRQAIL 87

```
Db 1 MDIDPYKFGATVELLSFLPSDFPSVRDLDTAAALFRALSPHCPTHTALROAIL 60
Qy 88 CWGELMTLATWGVNLEDPFRGDASRDLYVSVYNTMGLKFRQLMFIHISCLTFRGRTV 147
Db 61 CWGELMTLATWGVNLEDPFRGDASRDLYVSVYNTMGLKFRQLMFIHISCLTFRGRTV 115
Qy 148 IXYLVSGVWIRTPPAPRPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRR 207
Db 116 IXYLVSGVWIRTPPAPRPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRR 175
Qy 208 SQSREPOC 215
Db 176 TQSRBEOC 183
```

```
RESULT 6
US-10-274-616-6
; Sequence 6, Application US/10274616
; GENERAL INFORMATION:
; APPLICANT: Birkett, Ashley J.
; TITLE OF INVENTION: INFLUENZA IMMUNOGEN AND VACCINE
; FILE REFERENCE: ICC 127.0 4564/88545
; CURRENT APPLICATION NUMBER: US/10/274,616
; CURRENT FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: 09/930,915
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 10/080,299
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 6
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Spermophilus variegatus
US-10-274-616-6
```

```
Query Match 65.7%; Score 754.5; DB 6; Length 217;
Best Local Similarity 65.6%; Pred. No. 1,5e-47;
Matches 143; Conservative 21; Mismatches 43; Indels 11; Gaps 2;
```

```
Qy 3 ISSIFSRIDPTVOASKLCGLMGMNDIDPYKFGATVELLSFLPSDFPSVRDLDTAA 62
Db 6 LCLVACVCPVVOASKLCGLMGMNDIDPYKFGATVELLSFLPSDFPSVRDLDTAA 65
Qy 63 ALYREALSPHCSPHRTALROAILCWGELMTLATWGVNLEDPFRGDASRDLYVSVY 122
Db 66 ALYBELTGRHCSPHRTALROAILCWGELMTLATWGVNLEDPFRGDASRDLYVSVY 119
Qy 123 TMWGLKFRQLMFIHISCLTFRGRTVIEYVSGVWIRTPPAPRPNAPILSTLPETTV 182
Db 120 TMWGLKFRQLMFIHISCLTFRGRTVIEYVSGVWIRTPPAPRPNAPILSTLPETTV 179
Qy 183 RRGSRARSPRRRTSPRRRSQSPRRRSQSPRRRSQSPRRRSQSPRRRSQSPRRRS 215
Db 180 RRGSRARSPRRRTSPRRRSQSPRRRSQSPRRRSQSPRRRSQSPRRRSQSPRRRS 217
```

```
RESULT 7
PCT-US02-17536-17
; Sequence 17, Application PC/TUS0217536
; GENERAL INFORMATION:
; APPLICANT: Holoishitz, Joseph
; TITLE OF INVENTION: Methods and Compositions for the Treatment of Diseases Associated
; FILE REFERENCE: UM-07135
; CURRENT APPLICATION NUMBER: PCT/US02/17536
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 60/295,691
; PRIOR FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent in version 3.1
```

```
; SEQ ID NO 17
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
PCT-US02-17536-17
```

```
Query Match 62.9%; Score 722.5; DB 1; Length 169;
Best Local Similarity 87.7%; Pred. No. 2.6e-45;
Matches 142; Conservative 1; Mismatches 6; Indels 13; Gaps 2;
```

```
Qy 28 MDIDPYKFGATVELLSFLPSDFPSVRDLDTAAALFRALSPHCPTHTALROAIL 87
Db 1 MDIDPYKFGATVELLSFLPSDFPSVRDLDTAAALFRALSPHCPTHTALROAIL 60
Qy 88 CWGELMTLATWGVNLEDPFRGDASRDLYVSVYNTMGLKFRQLMFIHISCLTFRGRTV 134
Db 61 CWGELMTLATWGVNLEDPFRGDASRDLYVSVYNTMGLKFRQLMFIHISCLTFRGRTV 120
Qy 135 FHSICLTFGRRTVIEYVSGVWIRTPPAPRPNAPILSTLP 176
Db 121 FHSICLTFGRRTVIEYVSGVWIRTPPAPRPNAPILSTLP 162
```

```
RESULT 8
PCT-US02-17536-18
; Sequence 18, Application PC/TUS0217536
; GENERAL INFORMATION:
; APPLICANT: Holoishitz, Joseph
; TITLE OF INVENTION: Methods and Compositions for the Treatment of Diseases Associated
; FILE REFERENCE: UM-07135
; CURRENT APPLICATION NUMBER: PCT/US02/17536
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 60/295,691
; PRIOR FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 18
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
PCT-US02-17536-18
```

```
Query Match 62.9%; Score 722.5; DB 1; Length 169;
Best Local Similarity 87.0%; Pred. No. 2.6e-45;
Matches 141; Conservative 1; Mismatches 7; Indels 13; Gaps 2;
```

```
Qy 28 MDIDPYKFGATVELLSFLPSDFPSVRDLDTAAALFRALSPHCPTHTALROAIL 87
Db 1 MDIDPYKFGATVELLSFLPSDFPSVRDLDTAAALFRALSPHCPTHTALROAIL 60
Qy 88 CWGELMTLATWGVNLEDPFRGDASRDLYVSVYNTMGLKFRQLMFIHISCLTFRGRTV 134
Db 61 CWGELMTLATWGVNLEDPFRGDASRDLYVSVYNTMGLKFRQLMFIHISCLTFRGRTV 120
Qy 135 FHSICLTFGRRTVIEYVSGVWIRTPPAPRPNAPILSTLP 176
Db 121 FHSICLTFGRRTVIEYVSGVWIRTPPAPRPNAPILSTLP 162
```

```
RESULT 9
US-10-274-616-5
; Sequence 5, Application US/10274616
; GENERAL INFORMATION:
; APPLICANT: Birkett, Ashley J.
; TITLE OF INVENTION: INFLUENZA IMMUNOGEN AND VACCINE
; FILE REFERENCE: ICC 127.0 4564/88545
; CURRENT APPLICATION NUMBER: US/10/274,616
```

;; CURRENT FILING DATE: 2002-10-21
;; PRIOR APPLICATION NUMBER: 09/930,915
;; PRIOR FILING DATE: 2001-08-15
;; PRIOR APPLICATION NUMBER: 10/080,299
;; PRIOR FILING DATE: 2002-02-21
;; NUMBER OF SEQ ID NOS: 87
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 5
;; LENGTH: 183
;; TYPE: PRT
;; ORGANISM: Marmota monax
;; US-10-274-616-5

Query Match 57.6%; Score 661; DB 6; Length 183;
Best Local Similarity 66.3%; Pred. No. 7,8e-41;
Matches 124; Conservative 21; Mismatches 32; Indels 10; Gaps 2;

QY 28 MDIDPKFEGATVELISLPSPDFSVRLDLSALYREALSPHSCPHHTALROAIL 87
DB 1 MDIDPKFEGSSYOLNPLDFFPDNLALVDATALYEBELTGREHSPHHTALROALV 60
QY 88 CMGELMTLTATWGVNLEDEPFGDASRDIVSVYNTNMGAKEROLLMFHSICLTFGRRTV 147
DB 61 CMDELTKLJAMSSNITSQV-----KTIIVNHNVDTWGLKTRQSLMFTLSCLTFGQHTV 115
QY 148 IEVLSEFGWIRTPPAYRPPNAPILSTLEPTTVRRRG-----RSPRRRTSPRRRSQS 202
DB 116 QELVSEFGWIRTPPAYRPPNAPILSTLEPTTVRRRGCGARASRSPRRRTSPRRRSQS 175
QY 203 PRRRSQ 209
DB 176 PRRRSQ 182

RESULT 10
US-09-863-054-21
Sequence 21, Application US/09863054
GENERAL INFORMATION:
APPLICANT: Chisari, Francis V.
TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
Lymphocyte Responses to Hepatitis B Virus
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/863,054
FILING DATE: 21-May-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/935,898
FILING DATE: 26-AUG-1992
APPLICATION NUMBER: US 08/024,120
FILING DATE: 26-FEB-1993
APPLICATION NUMBER: US 08/396,283
FILING DATE: 27-FEB-1995
APPLICATION NUMBER: US 08/463,486
FILING DATE: 05-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Lockyer, Jean M.
REGISTRATION NUMBER: 44,879
REFERENCE/DOCKET NUMBER: 014740-000421US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200

;; TELEFAX: (415) 576-0300
;; INFORMATION FOR SEQ ID NO: 21:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 20 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: <Unknown>
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-09-863-054-21

Query Match 9.9%; Score 114; DB 5; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.044;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 152 VSEFGWIRTPPAYRPPNAPIL 171
DB 1 VSEFGWIRTPPAYRPPNAPIL 20

RESULT 11
US-10-218-140-1758
Sequence 1758, Application US/10218140
GENERAL INFORMATION:
APPLICANT: Shinketo, Richard A.
TITLE OF INVENTION: NOVEL POLYPEPTIDES AND POLYPEPTIDES
FILE REFERENCE: 15966-543 CON
CURRENT APPLICATION NUMBER: US/10/218,140
CURRENT FILING DATE: 2002-08-12
PRIOR APPLICATION NUMBER: 09/540,763
PRIOR FILING DATE: 2000-03-30
PRIOR APPLICATION NUMBER: 60/127,728
PRIOR FILING DATE: 1999-04-05
PRIOR APPLICATION NUMBER: 60/127,636
PRIOR FILING DATE: 1999-04-02
PRIOR APPLICATION NUMBER: 60/127,607
PRIOR FILING DATE: 1999-03-31
NUMBER OF SEQ ID NOS: 6322
SOFTWARE: Curator Version 1.0
SEQ ID NO 1758
LENGTH: 312
TYPE: PRT
ORGANISM: Homo sapiens
US-10-218-140-1758

Query Match 9.9%; Score 113.5; DB 6; Length 312;
Best Local Similarity 50.0%; Pred. No. 0.51;
Matches 28; Conservative 7; Mismatches 8; Indels 13; Gaps 3;

QY 165 RPPNAPIL-----STLEPTTVRRRGSPRRRTSPR--RRRSQPRRRSQRE 212
DB 222 KRRSPFLRYRRSRSP-----RRRSRSPKRRSPSPRRHRHSKSPRRRSRSD 272

RESULT 12
US-09-863-054-19
Sequence 19, Application US/09863054
GENERAL INFORMATION:
APPLICANT: Chisari, Francis V.
TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
Lymphocyte Responses to Hepatitis B Virus
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk


```
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/09/863,054
  FILING DATE: 21-May-2001
  CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 07/935,898
  FILING DATE: 26-AUG-1992
  APPLICATION NUMBER: US 08/024,120
  FILING DATE: 26-FEB-1993
  APPLICATION NUMBER: US 08/396,283
  FILING DATE: 27-FEB-1995
  APPLICATION NUMBER: US 08/463,486
  FILING DATE: 05-JUN-1995
ATTORNEY/AGENT INFORMATION:
  NAME: Lockyer, Jean M.
  REGISTRATION NUMBER: 44,879
  REFERENCE/DOCKET NUMBER: 014740-000421US
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (415) 576-0200
  TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 19:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 20 amino acids
    TYPE: amino acid
    STRANDEDNESS: <Unknown>
    TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-863-054-19

Query Match
Best Local Similarity 9.6%; Score 110; DB 5; Length 20;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 LMFHISCLTFRGTYIEYL 151
DB 1 LMFHISCLTFRGTYIEYL 20

RESULT 13
US-10-226-007-714
; Sequence 714, Application US/10226007
; GENERAL INFORMATION:
; APPLICANT: Myriad Genetics, Inc.
; APPLICANT: Morham, Scott
; APPLICANT: Zavitz, Kenton
; APPLICANT: Hobden, Adrian
; TITLE OF INVENTION: Compositions and Therapeutic Methods for Viral Infection
; FILE REFERENCE: 5005.01
; CURRENT APPLICATION NUMBER: US/10/226,007
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 60/313,883
; PRIOR FILING DATE: 2001-08-21
; NUMBER OF SEQ ID NOS: 1673
; SOFTWARE: Patent version 3.1
; SEQ ID NO 714
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Hepatitis B virus
US-10-226-007-714

Query Match
Best Local Similarity 9.6%; Score 110; DB 6; Length 20;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 149 EYLVSFGWIRTPPAYRPPN 168
DB 1 EYLVSFGWIRTPPAYRPPN 20
```

```
RESULT 14
US-10-226-007-646
; Sequence 646, Application US/10226007
; GENERAL INFORMATION:
; APPLICANT: Myriad Genetics, Inc.
; APPLICANT: Morham, Scott
; APPLICANT: Zavitz, Kenton
; APPLICANT: Hobden, Adrian
; TITLE OF INVENTION: Compositions and Therapeutic Methods for Viral Infection
; FILE REFERENCE: 5005.01
; CURRENT APPLICATION NUMBER: US/10/226,007
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 60/313,883
; PRIOR FILING DATE: 2001-08-21
; NUMBER OF SEQ ID NOS: 1673
; SOFTWARE: Patent version 3.1
; SEQ ID NO 646
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Hepatitis B virus
US-10-226-007-646

Query Match
Best Local Similarity 9.5%; Score 109; DB 6; Length 20;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 157 WIRTPPAYRPPNAPILSTLP 176
DB 1 WIRTPPAYRPPNAPILSTLP 20

RESULT 15
US-10-226-007-679
; Sequence 679, Application US/10226007
; GENERAL INFORMATION:
; APPLICANT: Myriad Genetics, Inc.
; APPLICANT: Morham, Scott
; APPLICANT: Zavitz, Kenton
; APPLICANT: Hobden, Adrian
; TITLE OF INVENTION: Compositions and Therapeutic Methods for Viral Infection
; FILE REFERENCE: 5005.01
; CURRENT APPLICATION NUMBER: US/10/226,007
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 60/313,883
; PRIOR FILING DATE: 2001-08-21
; NUMBER OF SEQ ID NOS: 1673
; SOFTWARE: Patent version 3.1
; SEQ ID NO 679
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Hepatitis B virus
US-10-226-007-679

Query Match
Best Local Similarity 9.5%; Score 109; DB 6; Length 20;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 154 FGWIRTPPAYRPPNAPILS 173
DB 1 FGWIRTPPAYRPPNAPILS 20

Search completed: February 3, 2003, 09:35:00
Job time : 17.6045 secs
```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 09:23:57 ; Search time 153.571 Seconds
(without alignments)
902.626 Million cell updates/sec

Title: US-09-890-752A-2

Perfect score: 1148
Sequence: 1 MFLSSIFSRHIGDPTVQASKL.....RRRSQSPRRRSQSRFFQC 215

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched:

4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents AA Main:*

1: /cgn2_6/ptodata/1/paa/PCRTUS_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep:*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep:*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep:*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep:*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep:*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pep:*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1148	100.0	215	22	US-09-890-752A-2
2	1048.5	91.3	212	21	US-09-791-537-92352
3	1048.5	91.3	212	22	US-09-848-616-114
4	1048.5	91.3	212	26	US-10-243-739-54
5	1048.5	91.3	212	26	US-10-244-065-54
6	1040.5	90.6	212	1	PCT-US97-15500-3

7	1040.5	90.6	212	21	US-09-705-547-11	Sequence 11, App1
8	1040.5	90.6	212	21	US-09-791-537-92352	Sequence 92352, A
9	1040.5	90.6	212	23	US-09-929-955-11	Sequence 11, App1
10	1040.5	90.6	212	25	US-10-104-966-11	Sequence 11, App1
11	1040.5	90.6	212	27	US-60-229-175-11	Sequence 11, App1
12	1038.5	90.5	212	21	US-09-791-537-125025	Sequence 125025, A
13	1037.5	90.4	212	21	US-09-718-095-12	Sequence 12, App1
14	1037.5	90.4	212	21	US-09-791-537-15944	Sequence 15944, A
15	1037.5	90.4	212	21	US-09-791-537-47294	Sequence 47294, A
16	1037.5	90.4	212	21	US-09-791-537-125027	Sequence 125027, A
17	1037.5	90.4	212	22	US-09-848-616-107	Sequence 107, App
18	1037.5	90.4	212	22	US-09-848-616-108	Sequence 108, App
19	1037.5	90.4	212	22	US-09-848-616-109	Sequence 109, App
20	1037.5	90.4	212	25	US-10-130-915-15	Sequence 15, App1
21	1037.5	90.4	212	26	US-10-243-739-47	Sequence 47, App1
22	1037.5	90.4	212	26	US-10-243-739-48	Sequence 48, App1
23	1037.5	90.4	212	26	US-10-244-065-47	Sequence 47, App1
24	1037.5	90.4	212	26	US-10-244-065-48	Sequence 48, App1
25	1035.5	90.2	212	21	US-09-791-537-125030	Sequence 125030, A
26	1035.5	90.2	212	21	US-09-791-537-125036	Sequence 125036, A
27	1035.5	90.2	212	21	US-09-848-616-112	Sequence 112, App
28	1035.5	90.2	212	22	US-09-848-616-113	Sequence 113, App
29	1035.5	90.2	212	26	US-10-243-739-52	Sequence 52, App1
30	1035.5	90.2	212	26	US-10-243-739-53	Sequence 53, App1
31	1035.5	90.2	212	26	US-10-244-065-52	Sequence 52, App1
32	1035.5	90.2	212	26	US-10-244-065-53	Sequence 53, App1
33	1034.5	90.1	212	21	US-09-791-537-125001	Sequence 53, App1
34	1034.5	90.1	212	21	US-09-791-537-125036	Sequence 125036, A
35	1034.5	90.1	212	21	US-09-791-537-125388	Sequence 125388, A
36	1034.5	90.1	212	22	US-09-848-616-115	Sequence 115, App
37	1034.5	90.1	212	22	US-09-848-616-116	Sequence 116, App
38	1034.5	90.1	212	22	US-09-848-616-117	Sequence 117, App
39	1034.5	90.1	212	26	US-10-243-739-49	Sequence 49, App1
40	1034.5	90.1	212	26	US-10-243-739-55	Sequence 55, App1
41	1034.5	90.1	212	26	US-10-243-739-56	Sequence 56, App1
42	1034.5	90.1	212	26	US-10-243-739-57	Sequence 57, App1
43	1034.5	90.1	212	26	US-10-244-065-49	Sequence 49, App1
44	1034.5	90.1	212	26	US-10-244-065-55	Sequence 55, App1
45	1034.5	90.1	212	26	US-10-244-065-56	Sequence 56, App1

ALIGNMENTS

RESULT 1
US-09-890-752A-2
Sequence 2, Application US/09890752A
GENERAL INFORMATION:
APPLICANT: Hilt, Eberhard
TITLE OF INVENTION: Particles for Gene Therapy
FILE REFERENCE: 107070-120 (VOS-013)
CURRENT APPLICATION NUMBER: US/09/890, 752A
PRIOR FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: PCT/DE00/00363
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: DE 199 04 800.2
NUMBER OF SEQ ID NOS: 21
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 215
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Fusion protein comprising a HBcAg, a cell-permeability-
US-09-890-752A-2

Query Match 100.0%; Score 1148; DB 22; Length 215;
Best Local Similarity 100.0%; Pred. No. 9.1e-110;
Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MPLSIFERIGDPTVQASKLCIGWLMGMDIDPYKEFGATVELLSFLPSDFPVSVDLDT 60
DB 1 MPLSIFERIGDPTVQASKLCIGWLMGMDIDPYKEFGATVELLSFLPSDFPVSVDLDT 60
QY 61 ASALYREALSEPERHCSPHHTALROAILCWGELMTLATWGVNLDPFRGDSRDVVS 120
DB 61 ASALYREALSEPERHCSPHHTALROAILCWGELMTLATWGVNLDPFRGDSRDVVS 120
QY 121 VNTNGLKFRQLMWHISCLTFGRETVIEYVSGWIRTPPAYRPPNAPILSTLPETT 180
DB 121 VNTNGLKFRQLMWHISCLTFGRETVIEYVSGWIRTPPAYRPPNAPILSTLPETT 180
QY 181 VRRGRSPRRRTSPRRRSOSRRRSQSRREPQC 215
DB 181 VRRGRSPRRRTSPRRRSOSRRRSQSRREPQC 215

RESULT 2
US-09-791-537-6815
; Sequence 6815, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biocomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patent version 3.0
; SEQ ID NO 6815
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Hepatitis B virus
US-09-791-537-6815

Query Match          91.3%; Score 1048.5; DB 21; Length 212;
Best Local Similarity 97.5%; Pred. No. 1.8e-99;
Matches 198; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 13 PTVOASKLCIGWLMGMDIDPYKEFGATVELLSFLPSDFPVSVDLDTASALYREALSE 72
DB 15 PTVOASKLCIGWLMGMDIDPYKEFGATVELLSFLPSDFPVSVDLDTASALYREALSE 74
QY 73 EHCSPHHTALROAILCWGELMTLATWGVNLDPFRGDSRDVVS YVNTNGLKFRQL 132
DB 75 EHCSPHHTALROAILCWGELMTLATWGVNLDPFRGDSRDVVS YVNTNGLKFRQL 129
QY 133 LWFHISCLTFGRETVIEYVSGWIRTPPAYRPPNAPILSTLPETT VRRGRSPRRRT 192
DB 130 LWFHISCLTFGRETVIEYVSGWIRTPPAYRPPNAPILSTLPETT VRRGRSPRRRT 189
QY 193 PSPRRRSOSRRRSQSRREPQC 215
DB 190 PSPRRRSOSRRRSQSRREPQC 212

RESULT 3
US-09-848-616-114
; Sequence 114, Application US/09848616
; GENERAL INFORMATION:
; APPLICANT: Seibel, Peter
; APPLICANT: Dunant, Nicolas
; APPLICANT: Bachmann, Martin
; APPLICANT: Tisoc, Alain
; APPLICANT: Lechner, Franziska
; TITLE OF INVENTION: Molecular Antigen Array
; FILE REFERENCE: 1700.0180002
; CURRENT APPLICATION NUMBER: US/09/848,616
; CURRENT FILING DATE: 2001-05-05
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: Patent version 2.1

```

```

; SEQ ID NO 114
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Hepatitis B virus
US-09-848-616-114

Query Match          91.3%; Score 1048.5; DB 22; Length 212;
Best Local Similarity 97.5%; Pred. No. 1.8e-99;
Matches 198; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 13 PTVOASKLCIGWLMGMDIDPYKEFGATVELLSFLPSDFPVSVDLDTASALYREALSE 72
DB 15 PTVOASKLCIGWLMGMDIDPYKEFGATVELLSFLPSDFPVSVDLDTASALYREALSE 74
QY 73 EHCSPHHTALROAILCWGELMTLATWGVNLDPFRGDSRDVVS YVNTNGLKFRQL 132
DB 75 EHCSPHHTALROAILCWGELMTLATWGVNLDPFRGDSRDVVS YVNTNGLKFRQL 129
QY 133 LWFHISCLTFGRETVIEYVSGWIRTPPAYRPPNAPILSTLPETT VRRGRSPRRRT 192
DB 130 LWFHISCLTFGRETVIEYVSGWIRTPPAYRPPNAPILSTLPETT VRRGRSPRRRT 189
QY 193 PSPRRRSOSRRRSQSRREPQC 215
DB 190 PSPRRRSOSRRRSQSRREPQC 212

RESULT 4
US-10-243-739-54
; Sequence 54, Application US/10243739
; GENERAL INFORMATION:
; APPLICANT: Bachmann, Martin F.
; APPLICANT: Seoni, Tazio
; APPLICANT: Lechner, Franziska
; TITLE OF INVENTION: In vivo Activation of Antigen Presenting Cells for Enhancement of
; FILE REFERENCE: 1700.0210001
; CURRENT APPLICATION NUMBER: US/10/243,739
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 60/318,967
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: Patent version 3.1
; SEQ ID NO 54
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Hepatitis B virus
US-10-243-739-54

Query Match          91.3%; Score 1048.5; DB 26; Length 212;
Best Local Similarity 97.5%; Pred. No. 1.8e-99;
Matches 198; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 13 PTVOASKLCIGWLMGMDIDPYKEFGATVELLSFLPSDFPVSVDLDTASALYREALSE 72
DB 15 PTVOASKLCIGWLMGMDIDPYKEFGATVELLSFLPSDFPVSVDLDTASALYREALSE 74
QY 73 EHCSPHHTALROAILCWGELMTLATWGVNLDPFRGDSRDVVS YVNTNGLKFRQL 132
DB 75 EHCSPHHTALROAILCWGELMTLATWGVNLDPFRGDSRDVVS YVNTNGLKFRQL 129
QY 133 LWFHISCLTFGRETVIEYVSGWIRTPPAYRPPNAPILSTLPETT VRRGRSPRRRT 192
DB 130 LWFHISCLTFGRETVIEYVSGWIRTPPAYRPPNAPILSTLPETT VRRGRSPRRRT 189
QY 193 PSPRRRSOSRRRSQSRREPQC 215
DB 190 PSPRRRSOSRRRSQSRREPQC 212

RESULT 5
US-10-244-065-54
; Sequence 54, Application US/10244065

```

```
/ GENERAL INFORMATION:
/ APPLICANT: Bachmann, Martin F.
/ APPLICANT: Storm, Tazio
/ APPLICANT: Maurer, Patrick
/ APPLICANT: Tisot, Alain
/ APPLICANT: Schwarz, Katrin
/ APPLICANT: Mejerink, Edwin
/ APPLICANT: Lipowsky, Gerd
/ APPLICANT: Rumpens, Paul
/ APPLICANT: Cielens, Indulis
/ APPLICANT: Renhofa, Regina
/ TITLE OF INVENTION: Packaging of Immunostimulatory Substances into Virus-like Particles
/ FILE REFERENCE: 1700.0220001
/ CURRENT APPLICATION NUMBER: US/10/244,065
/ PRIOR FILING DATE: 2002-09-16
/ PRIOR APPLICATION NUMBER: 60/374,145
/ PRIOR FILING DATE: 2002-04-22
/ PRIOR APPLICATION NUMBER: 60/318,994
/ PRIOR FILING DATE: 2001-09-14
/ NUMBER OF SEQ ID NOS: 73
/ SOFTWARE: Patent version 3.1
/ SEQ ID NO 54
/ LENGTH: 212
/ TYPE: PRT
/ ORGANISM: Hepatitis B virus
US-10-244-065-54

Query Match          91.3%; Score 1048.5; DB 26; Length 212;
Best Local Similarity 97.5%; Pred. No. 1.8e-99;
Matches 198; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 13 PTVOASKLCLGMLMGMDIDPYKFGATVELLSFLPSDFPSVRDLDTASALYREALBSP 72
DB 15 PTVOASKLCLGMLMGMDIDPYKFGATVELLSFLPSDFPSVRDLDTASALYREALBSP 74
QY 73 EHCSPHTALROAILCWGELMTLATWGVNLDEPFRGDASRDLYVSYVNTMGLKFRQL 132
DB 75 EHCSPHTALROAILCWGELMTLATWGVNLDEPFRGDASRDLYVSYVNTMGLKFRQL 129
QY 133 LMFHISCLTFGEETVIEIVSGWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRT 192
DB 130 LMFHISCLTFGEETVIEIVSGWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRT 189
QY 193 PPSRRRSQSPRRRSQSRBEOC 215
DB 190 PPSRRRSQSPRRRSQSRBEOC 212

RESULT 6
PCT-US97-15500-3
Sequence 3, Application PC/TUS9715500
GENERAL INFORMATION:
APPLICANT: The General Hospital Corporation
TITLE OF INVENTION: INHIBITION OF VIRAL REPLICATION
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/15500
FILING DATE: 03-SEP-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
```

```
/ APPLICATION NUMBER: 60/025,370
/ FILING DATE: 03-SEP-1996
/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Frazer, Janis K.
/ REGISTRATION NUMBER: 34,819
/ REFERENCE/DOCKET NUMBER: 08472/705W01
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617/542-5070
/ TELEFAX: 617/542-8906
/ TELEX: 200154
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 212 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
PCT-US97-15500-3

Query Match          90.6%; Score 1040.5; DB 1; Length 212;
Best Local Similarity 97.0%; Pred. No. 1.2e-98;
Matches 197; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

QY 13 PTVOASKLCLGMLMGMDIDPYKFGATVELLSFLPSDFPSVRDLDTASALYREALBSP 72
DB 15 PTVOASKLCLGMLMGMDIDPYKFGATVELLSFLPSDFPSVRDLDTASALYREALBSP 74
QY 73 EHCSPHTALROAILCWGELMTLATWGVNLDEPFRGDASRDLYVSYVNTMGLKFRQL 132
DB 75 EHCSPHTALROAILCWGELMTLATWGVNLDEPFRGDASRDLYVSYVNTMGLKFRQL 129
QY 133 LMFHISCLTFGEETVIEIVSGWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRT 192
DB 130 LMFHISCLTFGEETVIEIVSGWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRT 189
QY 193 PPSRRRSQSPRRRSQSRBEOC 215
DB 190 PPSRRRSQSPRRRSQSRBEOC 212

RESULT 7
US-09-705-547-11
Sequence 11, Application US/09705547
GENERAL INFORMATION:
APPLICANT: Catharina Halgren
TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
FILE REFERENCE: TRIPEP.023AUS
CURRENT APPLICATION NUMBER: US/09/705,547
CURRENT FILING DATE: 2000-11-03
PRIOR APPLICATION NUMBER: 60/229,175
PRIOR FILING DATE: 2000-08-29
NUMBER OF SEQ ID NOS: 15
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11
LENGTH: 212
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Hepatitis B virus C antigen and e antigen
OTHER INFORMATION: (HBCAg/HBeAg) sequence
US-09-705-547-11

Query Match          90.6%; Score 1040.5; DB 21; Length 212;
Best Local Similarity 97.0%; Pred. No. 1.2e-98;
Matches 197; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

QY 13 PTVOASKLCLGMLMGMDIDPYKFGATVELLSFLPSDFPSVRDLDTASALYREALBSP 72
DB 15 PTVOASKLCLGMLMGMDIDPYKFGATVELLSFLPSDFPSVRDLDTASALYREALBSP 74
QY 73 EHCSPHTALROAILCWGELMTLATWGVNLDEPFRGDASRDLYVSYVNTMGLKFRQL 132
```

```

Db 75 EHCSPHHTALROAILCWMGELMTLATWGVNLBDP-----ASRDLYVSYVNTMNGLKFRQL 129
Qy 133 LMFHISCLTGRGTVEIYLVSGVWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRT 192
Db 130 LMFHISCLTGRGTVEIYLVSGVWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRT 189
Qy 193 PSPRRRSQSPRRRSQSPRRSQC 215
Db 190 PSPRRRSQSPRRRSQSPRRSQC 212

```

RESULT 8
US-09-791-537-92352
Sequence 92352, Application US/09791537

```

GENERAL INFORMATION:
APPLICANT: Biomomix, Inc.
APPLICANT: Danzer, Joseph
APPLICANT: Debe, Derek
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
FILE REFERENCE: 261/210
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PatentIn version 3.0
SEQ ID NO 92352
LENGTH: 212
TYPE: PRT
ORGANISM: Hepatitis B virus
US-09-791-537-92352

```

Query Match 90.6%; Score 1040.5; DB 21; Length 212;
Best Local Similarity 97.0%; Pred. No. 1.2e-98;
Matches 197; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

```

Qy 13 PTVOASKLCIGWIMGMDIDPYKEFGATVELLSFLPSDFPSPVRDLDTASALYREALBSP 72
Db 15 PTVOASKLCIGWIMGMDIDPYKEFGATVELLSFLPSDFPSPVRDLDTASALYREALBSP 74
Qy 73 EHCSPHHTALROAILCWMGELMTLATWGVNLBDP-----ASRDLYVSYVNTMNGLKFRQL 132
Db 75 EHCSPHHTALROAILCWMGELMTLATWGVNLBDP-----ASRDLYVSYVNTMNGLKFRQL 129
Qy 133 LMFHISCLTGRGTVEIYLVSGVWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRT 192
Db 130 LMFHISCLTGRGTVEIYLVSGVWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRT 189
Qy 193 PSPRRRSQSPRRRSQSPRRSQC 215
Db 190 PSPRRRSQSPRRRSQSPRRSQC 212

```

RESULT 9
US-09-929-955-11
Sequence 11, Application US/09929955
GENERAL INFORMATION:
APPLICANT: Matti Salberg
APPLICANT: Catharina Hultgren
TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
FILE REFERENCE: TRIPEP.23AUS2
CURRENT FILING DATE: 2001-08-15
CURRENT APPLICATION NUMBER: US/09/929,955
PRIOR FILING DATE: 2000-11-03/705,547
PRIOR APPLICATION NUMBER: 60/229,175
PRIOR FILING DATE: 2000-08-29
PRIOR APPLICATION NUMBER: 60/225,767
PRIOR FILING DATE: 2000-08-17
NUMBER OF SEQ ID NOS: 49
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11

```

; LENGTH: 212
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis B virus C antigen and e antigen
; OTHER INFORMATION: (HbcAg/HbeAg) sequence
US-09-929-955-11

```

Query Match 90.6%; Score 1040.5; DB 23; Length 212;
Best Local Similarity 97.0%; Pred. No. 1.2e-98;
Matches 197; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

```

Qy 13 PTVOASKLCIGWIMGMDIDPYKEFGATVELLSFLPSDFPSPVRDLDTASALYREALBSP 72
Db 15 PTVOASKLCIGWIMGMDIDPYKEFGATVELLSFLPSDFPSPVRDLDTASALYREALBSP 74
Qy 73 EHCSPHHTALROAILCWMGELMTLATWGVNLBDP-----ASRDLYVSYVNTMNGLKFRQL 132
Db 75 EHCSPHHTALROAILCWMGELMTLATWGVNLBDP-----ASRDLYVSYVNTMNGLKFRQL 129
Qy 133 LMFHISCLTGRGTVEIYLVSGVWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRT 192
Db 130 LMFHISCLTGRGTVEIYLVSGVWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRT 189
Qy 193 PSPRRRSQSPRRRSQSPRRSQC 215
Db 190 PSPRRRSQSPRRRSQSPRRSQC 212

```

RESULT 10
US-10-104-966-11
Sequence 11, Application US/10104966

```

GENERAL INFORMATION:
APPLICANT: Matti Salberg
APPLICANT: Catharina Hultgren
TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
FILE REFERENCE: TRIPEP.23AUS2
CURRENT FILING DATE: 2002-03-22
PRIOR APPLICATION NUMBER: 09/705,547
PRIOR FILING DATE: 2000-11-03
PRIOR APPLICATION NUMBER: 60/229,175
PRIOR FILING DATE: 2000-08-29
NUMBER OF SEQ ID NOS: 15
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11
LENGTH: 212
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Hepatitis B virus C antigen and e antigen
OTHER INFORMATION: (HbcAg/HbeAg) sequence
US-10-104-966-11

```

Query Match 90.6%; Score 1040.5; DB 25; Length 212;
Best Local Similarity 97.0%; Pred. No. 1.2e-98;
Matches 197; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

```

Qy 13 PTVOASKLCIGWIMGMDIDPYKEFGATVELLSFLPSDFPSPVRDLDTASALYREALBSP 72
Db 15 PTVOASKLCIGWIMGMDIDPYKEFGATVELLSFLPSDFPSPVRDLDTASALYREALBSP 74
Qy 73 EHCSPHHTALROAILCWMGELMTLATWGVNLBDP-----ASRDLYVSYVNTMNGLKFRQL 132
Db 75 EHCSPHHTALROAILCWMGELMTLATWGVNLBDP-----ASRDLYVSYVNTMNGLKFRQL 129
Qy 133 LMFHISCLTGRGTVEIYLVSGVWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRT 192
Db 130 LMFHISCLTGRGTVEIYLVSGVWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRT 189
Qy 193 PSPRRRSQSPRRRSQSPRRSQC 215

```

Db 190 PPSRRRSQSPRRRSQSRSSQC 212

```
RESULT 11
US-60-229-175-11
; Sequence 11, Application US/60229175
; GENERAL INFORMATION:
; APPLICANT: Matci Saliberg
; APPLICANT: Catharina Holtegren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.023PR
; CURRENT APPLICATION NUMBER: US/60/229,175
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 212
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis B virus C antigen and e antigen
US-60-229-175-11
```

Query Match 90.6%; Score 1040.5; DB 27; Length 212,
Best Local Similarity 97.0%; Pred. No. 1,2e-98;
Matches 197; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

```
QY 13 PTVQASKLCIGMLMGMDIDPYKEFGATVELLSFLPSDFPSPVVDLDTASALYREALSP 72
DB 15 PTVQASKLCIGMLMGMDIDPYKEFGATVELLSFLPSDFPSPVVDLDTASALYREALSP 74
QY 73 EHCSPHHTALROAILCQGLMTLATWGVNLDPDFPRGASRLVVSYYNTMGLKFRQL 132
DB 75 EHCSPHHTALROAILCQGLMTLATWGVNLDP-----ASRLVVSYYNTMGLKFRQL 129
QY 133 LMFHISCLTFGRETVEIYVSGWIRTPPAYRPNAPILSTLPTTVVRRGRSPRRRT 192
DB 130 LMFHISCLTFGRETVEIYVSGWIRTPPAYRPNAPILSTLPTTVVRRGRSPRRRT 189
QY 193 PPSRRRSQSPRRRSQSRSSQC 215
DB 190 PPSRRRSQSPRRRSQSRSSQC 212
```

```
RESULT 12
US-09-791-537-125025
; Sequence 125025, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 125025
; LENGTH: 212
; TYPE: PRF
; ORGANISM: Hepatitis B virus
US-09-791-537-125025
```

Query Match 90.5%; Score 1038.5; DB 21; Length 212,
Best Local Similarity 96.6%; Pred. No. 1,9e-98;
Matches 196; Conservative 1; Mismatches 1; Indels 5; Gaps 1;

```
QY 13 PTVQASKLCIGMLMGMDIDPYKEFGATVELLSFLPSDFPSPVVDLDTASALYREALSP 72
DB 15 PTVQASKLCIGMLMGMDIDPYKEFGATVELLSFLPSDFPSPVVDLDTASALYREALSP 74
```

```
QY 73 EHCSPHHTALROAILCQGLMTLATWGVNLDPDFPRGASRLVVSYYNTMGLKFRQL 132
DB 75 EHCSPHHTALROAILCQGLMTLATWGVNLDP-----ASRLVVSYYNTMGLKFRQL 129
QY 133 LMFHISCLTFGRETVEIYVSGWIRTPPAYRPNAPILSTLPTTVVRRGRSPRRRT 192
DB 130 LMFHISCLTFGRETVEIYVSGWIRTPPAYRPNAPILSTLPTTVVRRGRSPRRRT 189
QY 193 PPSRRRSQSPRRRSQSRSSQC 215
DB 190 PPSRRRSQSPRRRSQSRSSQC 212
```

```
RESULT 13
US-09-718-095-12
; Sequence 12, Application US/09718095
; GENERAL INFORMATION:
; APPLICANT: STUYVER Lieven
; APPLICANT: VAN GREYT Caroline
; APPLICANT: DE GENDT Sija
; TITLE OF INVENTION: New HBV Sequences
; FILE REFERENCE: 2551-52
; CURRENT APPLICATION NUMBER: US/09/718,095
; CURRENT FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: EP99870252.6
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: US60/169,287
; NUMBER OF SEQ ID NOS: 169
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 212
; TYPE: PRF
; ORGANISM: Hepatitis B virus
US-09-718-095-12
```

Query Match 90.4%; Score 1037.5; DB 21; Length 212;
Best Local Similarity 96.6%; Pred. No. 2,4e-98;
Matches 196; Conservative 1; Mismatches 1; Indels 5; Gaps 1;

```
QY 13 PTVQASKLCIGMLMGMDIDPYKEFGATVELLSFLPSDFPSPVVDLDTASALYREALSP 72
DB 15 PTVQASKLCIGMLMGMDIDPYKEFGATVELLSFLPSDFPSPVVDLDTASALYREALSP 74
QY 73 EHCSPHHTALROAILCQGLMTLATWGVNLDPDFPRGASRLVVSYYNTMGLKFRQL 132
DB 75 EHCSPHHTALROAILCQGLMTLATWGVNLDP-----ASRLVVSYYNTMGLKFRQL 129
QY 133 LMFHISCLTFGRETVEIYVSGWIRTPPAYRPNAPILSTLPTTVVRRGRSPRRRT 192
DB 130 LMFHISCLTFGRETVEIYVSGWIRTPPAYRPNAPILSTLPTTVVRRGRSPRRRT 189
QY 193 PPSRRRSQSPRRRSQSRSSQC 215
DB 190 PPSRRRSQSPRRRSQSRSSQC 212
```

```
RESULT 14
US-09-791-537-15944
; Sequence 15944, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15944
```

LENGTH: 212
TYPE: PRT
ORGANISM: Hepatitis B virus
US-09-791-537-15944

Query Match 90.4%; Score 1037.5; DB 21; Length 212;
Best Local Similarity 96.6%; Pred. No. 2.4e-98;
Matches 196; Conservative 1; Mismatches 1; Indels 5; Gaps 1;

QY 13 PTVQASKLCIGMWMGMDIDPYKEGATVELLSFLPSDFPSVRDLDTASALYREALESP 72
DB 15 PTVQASKLCIGMWMGMDIDPYKEGATVELLSFLPSDFPSVRDLDTASALYREALESP 74
QY 73 EHCSPHHTALROAILCWGELMTLATWGVNLEDEPERGDASRDVSVYNTNMGKFEROL 132
DB 75 EHCSPHHTALROAILCWGELMTLATWGVNLEDEPERGDASRDVSVYNTNMGKFEROL 129
QY 133 LMFHISCLTFGRETIVIEYLVSGVWIRTPPAYRPNAPILSTLPTTVVRRGRSPRRRT 192
DB 130 LMFHISCLTFGRETIVIEYLVSGVWIRTPPAYRPNAPILSTLPTTVVRRGRSPRRRT 189
QY 193 PSPRRRSQSPRRRSQSPRRRSQSPRRRSQSPRRRSQSPRRRSQSPRRRSQSPRRRSQ 215
DB 190 PSPRRRSQSPRRRSQSPRRRSQSPRRRSQSPRRRSQSPRRRSQSPRRRSQSPRRRSQ 212

RESULT 15
US-09-791-537-47294
Sequence 47294, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Bionomix, Inc.
APPLICANT: Debe, Derek
APPLICANT: Danzer, Joseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBER
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: Patentin version 3.0
SEQ ID NO 47294
LENGTH: 212
TYPE: PRT
ORGANISM: Hepatitis B virus
US-09-791-537-47294

Query Match 90.4%; Score 1037.5; DB 21; Length 212;
Best Local Similarity 96.6%; Pred. No. 2.4e-98;
Matches 196; Conservative 1; Mismatches 1; Indels 5; Gaps 1;

QY 13 PTVQASKLCIGMWMGMDIDPYKEGATVELLSFLPSDFPSVRDLDTASALYREALESP 72
DB 15 PTVQASKLCIGMWMGMDIDPYKEGATVELLSFLPSDFPSVRDLDTASALYREALESP 74
QY 73 EHCSPHHTALROAILCWGELMTLATWGVNLEDEPERGDASRDVSVYNTNMGKFEROL 132
DB 75 EHCSPHHTALROAILCWGELMTLATWGVNLEDEPERGDASRDVSVYNTNMGKFEROL 129
QY 133 LMFHISCLTFGRETIVIEYLVSGVWIRTPPAYRPNAPILSTLPTTVVRRGRSPRRRT 192
DB 130 LMFHISCLTFGRETIVIEYLVSGVWIRTPPAYRPNAPILSTLPTTVVRRGRSPRRRT 189
QY 193 PSPRRRSQSPRRRSQSPRRRSQSPRRRSQSPRRRSQSPRRRSQSPRRRSQSPRRRSQ 215
DB 190 PSPRRRSQSPRRRSQSPRRRSQSPRRRSQSPRRRSQSPRRRSQSPRRRSQSPRRRSQ 212

Search completed: February 3, 2003, 09:34:08
Job time : 154.571 sec

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 09:25:53 ; Search time 7.86585 Seconds
(without alignments)
551.547 Million cell updates/sec

Title: US-09-890-752a-2

Sequence: 1 MRLSIFRIGDPTVQASKL.....RRRSQSPRRRSQSPRQC 215

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 122226 segs, 20178551 residues

Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCF_NEW_PUB pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/PCFUS_PUBCOMB pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1040.5	90.6	212	9	US-10-104-966-11
2	1040.5	90.6	212	10	US-09-829-955-11
3	977.5	85.1	212	9	US-10-209-264-4
4	938.5	81.6	183	10	US-09-812-862-12
5	933.5	81.3	397	10	US-09-812-862-6
6	928.5	80.9	724	9	US-10-068-059-12
7	928.5	80.9	746	9	US-10-068-059-6
8	926.5	80.7	183	10	US-09-816-230-4
9	922.5	80.4	185	9	US-10-068-059-2
10	922.5	80.4	185	10	US-09-816-230-5
11	920.5	80.2	289	10	US-09-812-862-8
12	863.5	75.2	351	10	US-09-812-862-4
13	778.5	67.8	690	9	US-10-068-059-10
14	772.5	67.3	709	9	US-10-068-059-8
15	771.5	67.2	170	9	US-10-068-059-4
16	754	65.7	152	10	US-09-816-230-6
17	736	64.1	152	10	US-09-816-230-7
18	607.5	52.9	346	10	US-09-812-862-7
19	157	13.7	420	10	US-09-812-862-10

20	114	9.9	20	10	US-09-839-447A-107	Sequence 107, App
21	114	9.9	23	10	US-09-839-447A-111	Sequence 111, App
22	112	9.8	20	10	US-09-839-447A-104	Sequence 104, App
23	111	9.7	20	9	US-09-466-035-84	Sequence 84, App1
24	111	9.7	20	10	US-09-839-447A-101	Sequence 101, App
25	110	9.6	20	10	US-09-839-447A-105	Sequence 105, App
26	109	9.5	20	10	US-09-839-447A-100	Sequence 100, App
27	108	9.4	20	10	US-09-839-447A-99	Sequence 99, App1
28	108	9.4	20	10	US-09-839-447A-106	Sequence 106, App1
29	105	9.1	20	10	US-09-839-447A-98	Sequence 98, App1
30	104	9.1	20	9	US-09-466-035-85	Sequence 85, App1
31	103	9.0	20	10	US-09-839-447A-108	Sequence 108, App
32	103	9.0	20	10	US-09-839-447A-110	Sequence 110, App
33	100	8.7	20	10	US-09-839-447A-97	Sequence 97, App1
34	100	8.7	20	10	US-09-839-447A-109	Sequence 109, App1
35	99	8.6	20	10	US-09-839-447A-96	Sequence 96, App1
36	99	8.6	20	10	US-09-839-447A-103	Sequence 103, App
37	94.5	8.2	376	10	US-09-825-299-1217	Sequence 102, App
38	92.5	8.1	20	10	US-09-839-447A-102	Sequence 109, App
39	85.5	7.4	123	10	US-09-825-299-1217	Sequence 121, App
40	83.5	7.3	144	10	US-09-825-299-1217	Sequence 113, App
41	83.5	7.3	219	10	US-09-894-018-113	Sequence 86, App1
42	83	7.2	15	9	US-09-466-035-86	Sequence 85, App1
43	81.5	7.1	266	10	US-09-864-761-37015	Sequence 37015, A
44	79.5	6.9	80	10	US-09-864-761-35898	Sequence 35898, A
45	79	6.9	2799	9	US-10-151-736-4	Sequence 4, App11

ALIGNMENTS

RESULT 1	
US-10-104-966-11	
Sequence 11, Application US/10104966	
Patent No. US2002015124A1	
GENERAL INFORMATION:	
APPLICANT: Macti Saliberg	
TITLE OF INVENTION: Vaccines containing Ribavirin and	
FILE REFERENCE: TRIPEP.23AUSCI	
CURRENT APPLICATION NUMBER: US/10/104, 966	
CURRENT FILING DATE: 2002-03-22	
PRIOR APPLICATION NUMBER: 09/705,547	
PRIOR FILING DATE: 2000-11-03	
PRIOR APPLICATION NUMBER: 60/229,175	
PRIOR FILING DATE: 2000-08-29	
NUMBER OF SEQ ID NOS: 15	
SOFTWARE: FastSeq for Windows Version 4.0	
SEQ ID NO 11	
LENGTH: 212	
TYPE: PRT	
ORGANISM: Artificial Sequence	
FEATURE:	
OTHER INFORMATION: Hepatitis B virus C antigen and e antigen	
OTHER INFORMATION: (HBCAg/HBeAg) sequence	
US-10-104-966-11	
Query Match	
Best Local Similarity 90.6%; Score 1040.5; DB 9; Length 212;	
Matches 197; Conservativity 0; Mismatches 1; Indels 5; Gaps 1;	
QY	13 PTQASKLCIGWLMGMDIDPKYKFGATVELLSFLPDPSPSVVDLDTASAYRELALSP 72
DB	15 PTQASKLCIGWLMGMDIDPKYKFGATVELLSFLPDPSPSVVDLDTASAYRELALSP 74
QY	73 EHCSPHTALROAILCIGWLMGMDIDPKYKFGATVELLSFLPDPSPSVVDLDTASAYRELALSP 132
DB	75 EHCSPHTALROAILCIGWLMGMDIDPKYKFGATVELLSFLPDPSPSVVDLDTASAYRELALSP 129
QY	133 LMFHISCLTFGRFTVEYVSRGVWTRTPAPRPAPILSTLPETTVVRRGRSRRRT 192
DB	130 LMFHISCLTFGRFTVEYVSRGVWTRTPAPRPAPILSTLPETTVVRRGRSRRRT 189

OY 193 PPRRRRSQSPRRRSQSGREPQC 215
Db 190 PPRRRRSQSPRRRSQSGRESQC 212

RESULT 2

US-09-929-955-11
Sequence 11, Application US/09929955
Patent No. US20020136740A1
GENERAL INFORMATION:
APPLICANT: Matti Salberg
APPLICANT: Catharina Hultgren
TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
FILE REFERENCE: TRIPEP.23AUS2
CURRENT APPLICATION NUMBER: US/09/929,955
PRIOR FILING DATE: 2001-08-15
PRIOR APPLICATION NUMBER: 09/705,547
PRIOR FILING DATE: 2000-11-03
PRIOR APPLICATION NUMBER: 60/229,175
PRIOR FILING DATE: 2000-08-29
PRIOR APPLICATION NUMBER: 60/225,767
PRIOR FILING DATE: 2000-08-17
NUMBER OF SEQ ID NOS: 49
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11
LENGTH: 212
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Hepatitis B virus C antigen and e antigen
US-09-929-955-11

Query Match 90.6%; Score 1040.5; DB 10; Length 212;

Best Local Similarity 97.0%; Pred. No. 2.4e-95;
Matches 197; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

OY 13 PTVQASKLCIGWLMGMDIDPYKFGATVELLSFLPSDFPSVVDLDTASALYREALESP 72
Db 15 PTVQASKLCIGWLMGMDIDPYKFGATVELLSFLPSDFPSVVDLDTASALYREALESP 74
OY 73 EHCSPHHTALRQAILCWEGLMTLATWGVNLEDFEFGDASRDLYSVYNTNGLKFRQL 132
Db 75 EHCSPHHTALRQAILCWEGLMTLATWGVNLEDFEFGDASRDLYSVYNTNGLKFRQL 129
OY 133 LMFHISCLTFGRETVIELVSGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRT 192
Db 130 LMFHISCLTFGRETVIELVSGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRT 189
OY 193 PPRRRRSQSPRRRSQSGREPQC 215
Db 190 PPRRRRSQSPRRRSQSGRESQC 212

RESULT 3

US-10-209-264-4
Sequence 4, Application US/10209264
Publication No. US2003000311A1
GENERAL INFORMATION:
APPLICANT: Oon, Chong Jin
Lhm, Gek Keow
Zhao, Yi
Chen, Wei Ning
TITLE OF INVENTION: A MUTANT HUMAN HEPATITIS B VIRAL STRAIN AND
USERS THEREOF
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Ladas & Parry
STREET: 26 West 61 Street
CITY: New York
STATE: New York

COUNTRY: USA

ZIP: 10023

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/209,264

FILING DATE: 31-Jul-2002

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/SG98/00046

FILING DATE: 19-Jan-1998

ATTORNEY/AGENT INFORMATION:
NAME: Maas, Clifford J.

REGISTRATION NUMBER: 30,086
REFERENCE/DOCKET NUMBER: U-013109-7

TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 708-1800

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
LENGTH: 212 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-10-209-264-4

Query Match 85.1%; Score 977.5; DB 9; Length 212;

Best Local Similarity 91.1%; Pred. No. 3.9e-89;
Matches 184; Conservative 5; Mismatches 8; Indels 5; Gaps 1;

OY 13 PTVQASKLCIGWLMGMDIDPYKFGATVELLSFLPSDFPSVVDLDTASALYREALESP 72
Db 15 PTVQASKLCIGWLMGMDIDPYKFGASAEILSLPSDFPSIRDLDTASALYREALESP 74
OY 73 EHCSPHHTALRQAILCWEGLMTLATWGVNLEDFEFGDASRDLYSVYNTNGLKFRQL 132
Db 75 EHCSPHHTALRQAILCWEGLMTLATWGVNLEDFEFGDASRDLYSVYNTNGLKFRQL 129
OY 133 LMFHISCLTFGRETVIELVSGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRT 192
Db 130 LMFHISCLTFGRETVIELVSGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRT 189
OY 193 PPRRRRSQSPRRRSQSGREPQC 214
Db 190 PPRRRRSQSPRRRSQSGRESQC 211

RESULT 4

US-09-812-862-12
Sequence 12, Application US/09812862
Patent No. US20020035081A1
GENERAL INFORMATION:
APPLICANT: Wanda, Jack R.
Scaglioni, Pier Paolo
Melegari, Margherita
TITLE OF INVENTION: INHIBITION OF HEPATITIS B REPLICATION
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/812,862
FILING DATE: 20-Mar-2001
CLASSIFICATION: <Unknown>
PRIORITY INFORMATION:
APPLICATION NUMBER: 08/667,073
FILING DATE: <Unknown>
APPLICATION NUMBER: 08/492,489
FILING DATE: 20-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/282001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 183 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-812-862-12

Query Match 81.8% Score 938.5; DB 10; Length 183;
Best Local Similarity 95.7%; Pred. No. 2,2e-85;
Matches 180; Conservative 0; Mismatches 3; Indels 5; Gaps 1;

QY 28 MDIDPYKEFGATVELLSFLPSDFPSPVRDLDTASALYREALSEPHCSPHHTALROAIL 87
DB 1 MDIDPYKEFGATVELLSFLPSDFPSPVRDLDTASALYREALSEPHCSPHHTALROAIL 60
QY 88 CWGELMTLATWGVNLDPDFRGSASRLDLYSVYNTNMGKFRQLMFWHISCLTFGRSTV 147
DB 61 CWGELMTLATWGVNLDPDFRGSASRLDLYSVYNTNMGKFRQLMFWHISCLTFGRSTV 115
QY 148 IEYLVSGVWIRTPPAYPPNAPILSTLPETTVRRRGSPRRTPSPRRRSQSPRRR 207
DB 116 IEYLVSGVWIRTPPAYPPNAPILSTLPETTVRRRGSPRRTPSPRRRSQSPRRR 175
QY 208 SOSREPOC 215
DB 176 SOSRESQC 183

RESULT 5
US-09-812-862-6
Sequence 6, Application US/09812862
Patent No. US20020035081A1
GENERAL INFORMATION:
APPLICANT: Wande, Jack R.
Seagalion, Pier Paolo
Margari, Margherita
TITLE OF INVENTION: INHIBITION OF HEPATITIS B REPLICATION
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/812,862
FILING DATE: 20-Mar-2001
CLASSIFICATION: <Unknown>
PRIORITY INFORMATION:

APPLICATION NUMBER: 08/667,073
FILING DATE: <Unknown>
APPLICATION NUMBER: 08/492,489
FILING DATE: 20-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/282001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 397 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-812-862-6

Query Match 81.3% Score 933.5; DB 10; Length 397;
Best Local Similarity 97.3%; Pred. No. 1.8e-84;
Matches 179; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 28 MDIDPYKEFGATVELLSFLPSDFPSPVRDLDTASALYREALSEPHCSPHHTALROAIL 87
DB 1 MDIDPYKEFGATVELLSFLPSDFPSPVRDLDTASALYREALSEPHCSPHHTALROAIL 60
QY 88 CWGELMTLATWGVNLDPDFRGSASRLDLYSVYNTNMGKFRQLMFWHISCLTFGRSTV 147
DB 61 CWGELMTLATWGVNLDPDFRGSASRLDLYSVYNTNMGKFRQLMFWHISCLTFGRSTV 115
QY 148 IEYLVSGVWIRTPPAYPPNAPILSTLPETTVRRRGSPRRTPSPRRRSQSPRRR 207
DB 116 IEYLVSGVWIRTPPAYPPNAPILSTLPETTVRRRGSPRRTPSPRRRSQSPRRR 175
QY 208 SOSR 211
DB 176 SOSR 179

RESULT 6
US-10-068-059-12
Sequence 12, Application US/10068059
Patent No. US20020155434A1
GENERAL INFORMATION:
APPLICANT: Mizzen, Lee A.
APPLICANT: Hongwei, Liu
TITLE OF INVENTION: HEPATITIS B VIRUS TREATMENT
FILE REFERENCE: 12071-017002
CURRENT APPLICATION NUMBER: US/10/068,059
CURRENT FILING DATE: 2002-06-04
PRIOR APPLICATION NUMBER: US 607266,733
PRIOR FILING DATE: 2001-02-05
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12
LENGTH: 724
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Fusion protein
US-10-068-059-12

Query Match 80.9% Score 928.5; DB 9; Length 724;
Best Local Similarity 94.2%; Pred. No. 1.2e-83;
Matches 179; Conservative 2; Mismatches 2; Indels 7; Gaps 2;
QY 28 MDIDPYKEFGATVELLSFLPSDFPSPVRDLDTASALYREALSEPHCSPHHTALROAIL 87
DB 1 MDIDPYKEFGATVELLSFLPSDFPSPVRDLDTASALYREALSEPHCSPHHTALROAIL 60

QY 88 CWGELMTLATVGVNLEDPFRGDASRDLYSVYNTNMGKFRQLMFIISCLTFGRRTV 147
 DB 61 CWGELMTLATVGVNLEDP-----ASRDLYSVYNTNMGKFRQLMFIISCLTFGRRTV 115
 QY 148 IEYLVSGVWIRTPPAYRPNNAPILSTLPETTVYRR--RGRSPRRRTSPRRRSQSPRR 205
 DB 116 IEYLVSGVWIRTPPAYRPNNAPILSTLPETTVYRRDRGRSPRRRTSPRRRSQSPRR 175
 QY 206 RRSQSRPQC 215
 DB 176 RRSQSRPQC 185

RESULT 7

US-10-068-059-6
 ; Sequence 6, Application US/10068059
 ; Patent No. US20020155434A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mizzzen, Lee A.
 ; APPLICANT: Hongwei, Liu
 ; APPLICANT: Siegel, Marvin
 ; TITLE OF INVENTION: HEPATITIS B VIRUS TREATMENT
 ; FILE REFERENCE: 12071-011002
 ; CURRENT APPLICATION NUMBER: US/10/068,059
 ; CURRENT FILING DATE: 2002-06-04
 ; PRIOR APPLICATION NUMBER: US 60/266,733
 ; PRIOR FILING DATE: 2001-02-05
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 6
 ; LENGTH: 746
 ; TYPE: PRP
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Fusion protein
 ; US-10-068-059-6

Query Match 80.9%; Score 928.5; DB 9; Length 746;
 Best Local Similarity 94.2%; Pred. No. 1,3e-83;
 Matches 179; Conservative 2; Mismatches 2; Indels 7; Gaps 2;

QY 28 MDIDPYKEFGATVELLSFLPSDFPFSVRDLDTASALYREALSPHCSPHHTALRQAIL 87
 DB 21 MDIDPYKEFGATVELLSFLPSDFPFSVRDLDTASALYREALSPHCSPHHTALRQAIL 80
 QY 88 CWGELMTLATVGVNLEDPFRGDASRDLYSVYNTNMGKFRQLMFIISCLTFGRRTV 147
 DB 81 CWGELMTLATVGVNLEDP-----ASRDLYSVYNTNMGKFRQLMFIISCLTFGRRTV 135
 QY 148 IEYLVSGVWIRTPPAYRPNNAPILSTLPETTVYRR--RGRSPRRRTSPRRRSQSPRR 205
 DB 136 IEYLVSGVWIRTPPAYRPNNAPILSTLPETTVYRRDRGRSPRRRTSPRRRSQSPRR 195
 QY 206 RRSQSRPQC 215
 DB 196 RRSQSRPQC 205

RESULT 8

US-09-916-230-4
 ; Sequence 4, Application US/09916230
 ; Patent No. US20020146422A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bachmann, Martin F.
 ; APPLICANT: Rennert, Wolfgang A.
 ; TITLE OF INVENTION: Compositions for Inducing Self-Specific Anti-IgE
 ; FILE REFERENCE: 1700,0140001
 ; CURRENT APPLICATION NUMBER: US/09/916,230
 ; CURRENT FILING DATE: 2001-07-27
 ; PRIOR APPLICATION NUMBER: US 60/221,841
 ; PRIOR FILING DATE: 2000-07-28

; NUMBER OF SEQ ID NOS: 35
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 4
 ; LENGTH: 183
 ; TYPE: PRP
 ; ORGANISM: Hepatitis B virus
 ; US-09-916-230-4

Query Match 80.7%; Score 926.5; DB 10; Length 183;
 Best Local Similarity 94.7%; Pred. No. 3.4e-84;
 Matches 178; Conservative 0; Mismatches 5; Indels 5; Gaps 1;

QY 28 MDIDPYKEFGATVELLSFLPSDFPFSVRDLDTASALYREALSPHCSPHHTALRQAIL 87
 DB 1 MDIDPYKEFGATVELLSFLPSDFPFSVRDLDTASALYREALSPHCSPHHTALRQAIL 60
 QY 88 CWGELMTLATVGVNLEDPFRGDASRDLYSVYNTNMGKFRQLMFIISCLTFGRRTV 147
 DB 61 CWGELMTLATVGVNLEDP-----ISRDLYSVYNTNMGKFRQLMFIISCLTFGRRTV 115
 QY 148 IEYLVSGVWIRTPPAYRPNNAPILSTLPETTVYRRDRGRSPRRRTSPRRRSQSPRR 207
 DB 116 IEYLVSGVWIRTPPAYRPNNAPILSTLPETTVYRRDRGRSPRRRTSPRRRSQSPRR 175
 QY 208 RRSQSRPQC 215
 DB 176 RRSQSRPQC 183

RESULT 9

US-10-068-059-2
 ; Sequence 2, Application US/10068059
 ; Patent No. US20020155434A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mizzzen, Lee A.
 ; APPLICANT: Hongwei, Liu
 ; APPLICANT: Siegel, Marvin
 ; TITLE OF INVENTION: HEPATITIS B VIRUS TREATMENT
 ; FILE REFERENCE: 12071-011002
 ; CURRENT APPLICATION NUMBER: US/10/068,059
 ; CURRENT FILING DATE: 2002-06-04
 ; PRIOR APPLICATION NUMBER: US 60/266,733
 ; PRIOR FILING DATE: 2001-02-05
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 185
 ; TYPE: PRP
 ; ORGANISM: Hepatitis B virus
 ; US-10-068-059-2

Query Match 80.4%; Score 922.5; DB 9; Length 185;
 Best Local Similarity 93.7%; Pred. No. 8.6e-84;
 Matches 178; Conservative 2; Mismatches 3; Indels 7; Gaps 2;

QY 28 MDIDPYKEFGATVELLSFLPSDFPFSVRDLDTASALYREALSPHCSPHHTALRQAIL 87
 DB 1 MDIDPYKEFGATVELLSFLPSDFPFSVRDLDTASALYREALSPHCSPHHTALRQAIL 60
 QY 88 CWGELMTLATVGVNLEDPFRGDASRDLYSVYNTNMGKFRQLMFIISCLTFGRRTV 147
 DB 61 CWGELMTLATVGVNLEDP-----ASRDLYSVYNTNMGKFRQLMFIISCLTFGRRTV 115
 QY 148 IEYLVSGVWIRTPPAYRPNNAPILSTLPETTVYRR--RGRSPRRRTSPRRRSQSPRR 205
 DB 116 IEYLVSGVWIRTPPAYRPNNAPILSTLPETTVYRRDRGRSPRRRTSPRRRSQSPRR 175
 QY 206 RRSQSRPQC 215
 DB 176 RRSQSRPQC 185

RESULT 10

US-09-916-230-5
Sequence 5, Application US/09916230
Patent No. US20020146422A1
GENERAL INFORMATION:
APPLICANT: Bachmann, Martin F.
APPLICANT: Renner, Wolfgang A.
TITLE OF INVENTION: Compositions for inducing Self-Specific Anti-IgE
TITLE OF INVENTION: Antibodies and uses thereof
FILE REFERENCE: 1700.0140001
CURRENT APPLICATION NUMBER: US/09/916,230
CURRENT FILING DATE: 2001-07-27
PRIOR APPLICATION NUMBER: US 60/221,841
PRIOR FILING DATE: 2000-07-28
NUMBER OF SEQ ID NOS: 35
SOFTWARE: Patentin version 3.0
SEQ ID NO 5
LENGTH: 185
TYPE: PRP
ORGANISM: Hepatitis B virus
US-09-916-230-5

Query Match 80.4%; Score 922.5; DB 10; Length 185;
Best Local Similarity 93.7%; Pred. No. 8,6e-84;
Matches 178; Conservative 2; Mismatches 3; Indels 7; Gaps 2;

QY 28 MDIDPKFGATVELLSRLPSDFPSVRLDLDTASALYREALSPHSCPHHTALRQAIL 87
DB 1 MDIDPKFGATVELLSRLPSDFPSVRLDLDTASALYREALSPHSCPHHTALRQAIL 60
QY 88 CMGELMTLATWGVNLEDPFRGDASRDVSVYNTNMGKLRQLMFWHISCLTFGRRTV 147
DB 61 CMGELMTLATWGVNLEDPFRGDASRDVSVYNTNMGKLRQLMFWHISCLTFGRRTV 115
QY 148 IEYVSGFWIRTPPAYRPPNAPILSTLPTTVRRGRSPRRRTSPRRRSQSPRR 205
DB 116 IEYVSGFWIRTPPAYRPPNAPILSTLPTTVRRGRSPRRRTSPRRRSQSPRR 175
QY 206 RRSQSPRRRTSPRRRSQSPRR 215
DB 176 RRSQSPRRRTSPRRRSQSPRR 185

RESULT 11
US-09-812-862-8
Sequence 8, Application US/09812862
Patent No. US20020035081A1
GENERAL INFORMATION:
APPLICANT: Wands, Jack R.
APPLICANT: Scaglioni, Pier Paolo
APPLICANT: Melegari, Margherita
TITLE OF INVENTION: INHIBITION OF HEPATITIS B REPLICATION
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSER: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/812,862
FILING DATE: 20-Mar-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/667,073
FILING DATE: <Unknown>
APPLICATION NUMBER: 08/492,489
FILING DATE: 20-JUN-1995

ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/282001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 289 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULAR TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-812-862-8

Query Match 80.2%; Score 920.5; DB 10; Length 289;
Best Local Similarity 96.2%; Pred. No. 2,4e-83;
Matches 176; Conservative 1; Mismatches 1; Indels 5; Gaps 1;

QY 28 MDIDPKFGATVELLSRLPSDFPSVRLDLDTASALYREALSPHSCPHHTALRQAIL 87
DB 1 MDIDPKFGATVELLSRLPSDFPSVRLDLDTASALYREALSPHSCPHHTALRQAIL 60
QY 88 CMGELMTLATWGVNLEDPFRGDASRDVSVYNTNMGKLRQLMFWHISCLTFGRRTV 147
DB 61 CMGELMTLATWGVNLEDPFRGDASRDVSVYNTNMGKLRQLMFWHISCLTFGRRTV 115
QY 148 IEYVSGFWIRTPPAYRPPNAPILSTLPTTVRRGRSPRRRTSPRRRSQSPRR 207
DB 116 IEYVSGFWIRTPPAYRPPNAPILSTLPTTVRRGRSPRRRTSPRRRSQSPRR 175
QY 208 SQS 210
DB 176 SST 178

RESULT 12
US-09-812-862-4
Sequence 4, Application US/09812862
Patent No. US20020035081A1
GENERAL INFORMATION:
APPLICANT: Wands, Jack R.
APPLICANT: Scaglioni, Pier Paolo
APPLICANT: Melegari, Margherita
TITLE OF INVENTION: INHIBITION OF HEPATITIS B REPLICATION
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSER: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/812,862
FILING DATE: 20-Mar-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/667,073
FILING DATE: <Unknown>
APPLICATION NUMBER: 08/492,489
FILING DATE: 20-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/282001

```
TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617/542-5070
/ TELEFAX: 617/542-8906
/ TELE: 200154
/ INFORMATION FOR SEQ ID NO: 4:
/ SEQUENCE CHARACTERISTICS:
/   LENGTH: 351 amino acids
/   TYPE: amino acid
/   TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-812-862-4

Query Match      75.2%; Score 863.5; DB 10; Length 351;
Best Local Similarity 87.5%; Pred. No. 1.3e-77;
Matches 168; Conservative 3; Mismatches 12; Indels 9; Gaps 2;

QY 28 MDIDPYKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCHSPHHTALRQAIL 87
DB 4 MDIDPYKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCHSPHHTALRQAIL 63
QY 88 CWGELMTLATWGVNLDEDFRGSARDLVSVYNTNMGKFRQLMFIHISCLTFGRETV 147
DB 64 CWGELMTLATWGVNLDEDFRGSARDLVSVYNTNMGKFRQLMFIHISCLTFGRETV 118
QY 148 IEYLVSGFWIRTPPAYRPPNAPILSTLPETTVRRR-----RSPPRRRSPRRRRSOSP 203
DB 119 IEYLVSGFWIRTPPAYRPPNAPILSTLPETTVIRRGARASRSPRRTPSPRRRRSONS 178
QY 204 RRRRSQSREPOC 215
DB 179 QSPTSNHSPTSC 190

RESULT 13
US-10-068-059-10
/ Sequence 10, Application US/10068059
/ Patent No. US20020155434A1
/ GENERAL INFORMATION:
/   APPLICANT: Mizzzen, Lee A.
/   APPLICANT: Hongwei, Liu
/   APPLICANT: Siegel, Marvin
/   TITLE OF INVENTION: HEPATITIS B VIRUS TREATMENT
/   FILE REFERENCE: 12071-017002
/   CURRENT APPLICATION NUMBER: US/10/068,059
/   CURRENT FILING DATE: 2002-06-04
/   PRIOR APPLICATION NUMBER: US 60/266,733
/   PRIOR FILING DATE: 2001-02-05
/   NUMBER OF SEQ ID NOS: 12
/   SOFTWARE: FastSeq for Windows Version 4.0
/   SEQ ID NO 10
/   LENGTH: 690
/   TYPE: PRT
/   ORGANISM: Artificial Sequence
/   FEATURE:
/   OTHER INFORMATION: Fusion protein
US-10-068-059-10

Query Match      67.8%; Score 778.5; DB 9; Length 690;
Best Local Similarity 94.9%; Pred. No. 6.9e-69;
Matches 148; Conservative 2; Mismatches 1; Indels 5; Gaps 1;

QY 28 MDIDPYKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCHSPHHTALRQAIL 87
DB 1 MDIDPYKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCHSPHHTALRQAIL 60
QY 88 CWGELMTLATWGVNLDEDFRGSARDLVSVYNTNMGKFRQLMFIHISCLTFGRETV 147
DB 61 CWGELMTLATWGVNLDEDFRGSARDLVSVYNTNMGKFRQLMFIHISCLTFGRETV 115
QY 148 IEYLVSGFWIRTPPAYRPPNAPILSTLPETTVRR 183
DB 116 IEYLVSGFWIRTPPAYRPPNAPILSTLPETTVRR 151
```

```
RESULT 14
US-10-068-059-8
/ Sequence 8, Application US/10068059
/ Patent No. US20020155434A1
/ GENERAL INFORMATION:
/   APPLICANT: Mizzzen, Lee A.
/   APPLICANT: Hongwei, Liu
/   APPLICANT: Siegel, Marvin
/   TITLE OF INVENTION: HEPATITIS B VIRUS TREATMENT
/   FILE REFERENCE: 12071-017002
/   CURRENT APPLICATION NUMBER: US/10/068,059
/   CURRENT FILING DATE: 2002-06-04
/   PRIOR APPLICATION NUMBER: US 60/266,733
/   PRIOR FILING DATE: 2001-02-05
/   NUMBER OF SEQ ID NOS: 12
/   SOFTWARE: FastSeq for Windows Version 4.0
/   SEQ ID NO 8
/   LENGTH: 709
/   TYPE: PRT
/   ORGANISM: Artificial Sequence
/   FEATURE:
/   OTHER INFORMATION: Fusion protein
US-10-068-059-8

Query Match      67.3%; Score 772.5; DB 9; Length 709;
Best Local Similarity 93.6%; Pred. No. 2.8e-68;
Matches 147; Conservative 2; Mismatches 3; Indels 5; Gaps 1;

QY 28 MDIDPYKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCHSPHHTALRQAIL 87
DB 21 MDIDPYKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCHSPHHTALRQAIL 80
QY 88 CWGELMTLATWGVNLDEDFRGSARDLVSVYNTNMGKFRQLMFIHISCLTFGRETV 147
DB 81 CWGELMTLATWGVNLDEDFRGSARDLVSVYNTNMGKFRQLMFIHISCLTFGRETV 135
QY 148 IEYLVSGFWIRTPPAYRPPNAPILSTLPETTVRR 184
DB 136 IEYLVSGFWIRTPPAYRPPNAPILSTLPETTVNAK 172

RESULT 15
US-10-068-059-4
/ Sequence 4, Application US/10068059
/ Patent No. US20020155434A1
/ GENERAL INFORMATION:
/   APPLICANT: Mizzzen, Lee A.
/   APPLICANT: Hongwei, Liu
/   APPLICANT: Siegel, Marvin
/   TITLE OF INVENTION: HEPATITIS B VIRUS TREATMENT
/   FILE REFERENCE: 12071-017002
/   CURRENT APPLICATION NUMBER: US/10/068,059
/   CURRENT FILING DATE: 2002-06-04
/   PRIOR APPLICATION NUMBER: US 60/266,733
/   PRIOR FILING DATE: 2001-02-05
/   NUMBER OF SEQ ID NOS: 12
/   SOFTWARE: FastSeq for Windows Version 4.0
/   SEQ ID NO 4
/   LENGTH: 170
/   TYPE: PRT
/   ORGANISM: Artificial Sequence
/   FEATURE:
/   OTHER INFORMATION: Fusion protein
US-10-068-059-4

Query Match      67.2%; Score 771.5; DB 9; Length 170;
Best Local Similarity 95.5%; Pred. No. 5.9e-69;
Matches 147; Conservative 1; Mismatches 1; Indels 5; Gaps 1;

QY 28 MDIDPYKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCHSPHHTALRQAIL 87
```

Db 21 MDIDPYKEGATVHELSEFLPSDFPSPYRDLLDTASALYREALSPHCSPHHTALROAIL 80
QY 88 CMGELMTLATWGVNLEDPFPRGDASRDLVSVYNTMGLKFRQLWFEHISCLTFGRETV 147
Db 81 CMGELMTLATWGVNLEDP-----ASRDLVSVYNTMGLKFRQLWFEHISCLTFGRETV 135
QY 148 IXYLVSRGVWIKRTPPAYRPPNAPILSTLPBTTVV 181
Db 136 LXYLVSRGVWIKRTPPAYRPPNAPILSTLPBTTVV 169

Search completed: February 3, 2003, 09:35:29
Job time : 9.86585 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 09:22:17 / Search time 13.1098 Seconds

(Without alignments)
482.536 Million cell updates/sec

Title: US-09-890-752A-2

Perfect score: 1148

Sequence: 1 MFLSSIFSRIGDPTVQASKL.....RRRSQSPRRRSQSPRRQPC 215

Scoring table:

BLOSUM62

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1040.5	90.6	212	3	US-08-968-747-3
2	1015.5	88.5	194	3	US-08-968-747-18
3	1014	88.3	199	3	US-08-968-747-21
4	1000.5	87.2	193	3	US-08-968-747-2
5	972	84.7	346	1	US-08-105-483-217
6	972	84.7	346	1	US-08-709-209-217
7	972	84.7	346	1	US-08-458-101-217
8	954.5	83.1	211	6	5196194-13
9	951.5	82.9	183	4	US-09-248-588-2
10	938.5	81.8	183	5	PCT-US96-10602-12
11	936.5	81.6	183	3	US-08-968-747-20
12	933.5	81.3	397	5	PCT-US96-10602-6
13	922.5	80.4	185	1	US-07-739-642-2
14	922.5	80.4	185	1	US-07-739-642-2
15	922.5	80.4	185	1	US-07-739-642-2
16	920.5	80.2	289	5	PCT-US96-10602-8
17	919.5	80.1	185	4	US-09-248-588-6
18	918.5	80.0	185	1	US-07-739-642-4
19	918.5	80.0	185	1	US-07-739-643-4
20	918.5	80.0	185	1	US-07-739-643-4
21	918.5	80.0	185	1	US-07-739-642-10
22	910.5	79.3	185	1	US-07-739-643-10
23	910.5	79.3	185	1	US-07-739-643-10
24	910.5	79.3	185	1	US-07-739-642-8
25	907.5	79.1	185	1	US-07-739-643-8
26	907.5	79.1	185	1	US-07-739-643-8
27	907.5	79.1	185	1	US-07-739-142-8

28	863.5	75.2	351	5	PCT-US96-10602-4	Sequence 4, Appli
29	844.5	73.6	159	4	US-08-445-585-3	Sequence 3, Appli
30	821.5	71.6	154	3	US-08-968-747-1	Sequence 1, Appli
31	821.5	71.6	155	3	US-08-968-747-17	Sequence 17, Appli
32	821.5	71.6	161	3	US-08-968-747-19	Sequence 19, Appli
33	772.5	67.3	214	1	US-07-739-642-12	Sequence 12, Appli
34	772.5	67.3	214	1	US-07-739-643-12	Sequence 12, Appli
35	772.5	67.3	214	1	US-07-739-642-12	Sequence 12, Appli
36	767.5	66.9	214	1	US-07-739-642-6	Sequence 6, Appli
37	767.5	66.9	214	1	US-07-739-642-6	Sequence 6, Appli
38	767.5	66.9	214	1	US-07-739-642-6	Sequence 6, Appli
39	757.5	66.0	217	4	US-09-248-588-9	Sequence 9, Appli
40	671	58.4	188	4	US-09-248-588-9	Sequence 9, Appli
41	607.5	52.9	346	5	PCT-US96-10602-2	Sequence 2, Appli
42	194	16.9	39	3	US-08-968-747-5	Sequence 5, Appli
43	175	15.2	305	4	US-09-248-588-11	Sequence 11, Appli
44	166	14.5	305	4	US-09-248-588-13	Sequence 13, Appli
45	157	13.7	420	5	PCT-US96-10602-10	Sequence 10, Appli

ALIGNMENTS

RESULT 1
US-08-968-747-3

Sequence 3, Application US/08968747

Patent No. 6060595

GENERAL INFORMATION:

APPLICANT: Scaglion et al.

TITLE OF INVENTION: INHIBITION OF VIRAL REPLICATION

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSER: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: Windows

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/968,747

FILING DATE: 03-SEP-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Fraser, Janis K.

REGISTRATION NUMBER: 34,819

REFERENCE/DOCKET NUMBER: 08472/705001

TELEPHONE/COMMUNICATION INFORMATION:

TELEPHONE: 617/542-5070

TELEFAX: 617/542-8906

TELEX: 200154

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 212 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULAR TYPE: protein

US-08-968-747-3

Query Match

Best local Similarity 97.0%

Matches 197; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

DB

13 PTVQASKLCLGIMGNDIDPYKFGATVELSLPSDFPSYRDLDTPASALYREALASP 72

15 PTVQASKLCLGIMGNDIDPYKFGATVELSLPSDFPSYRDLDTPASALYREALASP 74

QY 73 BHCSPHHTLRLQALLCWGSLMTLATVGVNLEDPFRGASDVLVSYVNTMGKLFROL 132

Thu Feb 6 14:44:24 2003

us-09-890-752a-2.raf

Page 2

Db 75 EHCSPHTALROAILCWGELMTLATWGVNLEDP-----ASRDVSVYNTNMGLKFRQL 129
Qy 133 LMFHTISCLTFGRVETIYLVSGFWIRTPPAYRPNAPILSTLPTTIVRRGRSPRR 192
Db 130 LMFHTISCLTFGRVETIYLVSGFWIRTPPAYRPNAPILSTLPTTIVRRGRSPRR 189
Qy 193 PSPRRRSQSPRRRSQSPRR 215
Db 190 PSPRRRSQSPRRRSQSPRR 212

RESULT 2

US-08-966-747-18
Sequence 18, Application US/08968747
Patent No. 6060595
GENERAL INFORMATION:
APPLICANT: Scaglioni et al.
TITLE OF INVENTION: INHIBITION OF VIRAL REPLICATION
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/968,747
FILING DATE: 03-SEP-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 08472/705001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 194 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-968-747-18

Query Match 88.5%; Score 1015.5; DB 3; Length 194;
Best Local Similarity 97.0%; Pred. No. 1.1e-103;
Matches 192; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

Qy 18 SKCLGMLMGMDIDPYKFGATVVELSFLPSDFPVSVDLDTASALYREALSPHCSP 77
Db 2 SKCLGMLMGMDIDPYKFGATVVELSFLPSDFPVSVDLDTASALYREALSPHCSP 61
Qy 78 HTALROAILCWGELMTLATWGVNLEDPFRGDASRDVSVYNTNMGLKFRQLMFHI 137
Db 62 HTALROAILCWGELMTLATWGVNLEDP-----ASRDVSVYNTNMGLKFRQLMFHI 116
Qy 138 SCLTFGRVETIYLVSGFWIRTPPAYRPNAPILSTLPTTIVRRGRSPRRTPSPRR 197
Db 117 SCLTFGRVETIYLVSGFWIRTPPAYRPNAPILSTLPTTIVRRGRSPRRTPSPRR 176
Qy 198 RRSQSPRRRSQSPRR 215
Db 177 RRSQSPRRRSQSPRR 194

RESULT 3

US-08-968-747-21
Sequence 21, Application US/08968747
Patent No. 6060595
GENERAL INFORMATION:
APPLICANT: Scaglioni et al.
TITLE OF INVENTION: INHIBITION OF VIRAL REPLICATION
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/968,747
FILING DATE: 03-SEP-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 08472/705001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 199 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-968-747-21

Query Match 88.3%; Score 1014; DB 3; Length 199;
Best Local Similarity 96.5%; Pred. No. 1.6e-103;
Matches 191; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 18 SKCLGMLMGMDIDPYKFGATVVELSFLPSDFPVSVDLDTASALYREALSPHCSP 77
Db 2 SKCLGMLMGMDIDPYKFGATVVELSFLPSDFPVSVDLDTASALYREALSPHCSP 61
Qy 78 HTALROAILCWGELMTLATWGVNLEDPFRGDASRDVSVYNTNMGLKFRQLMFHI 137
Db 62 HTALROAILCWGELMTLATWGVNLEDPFRGDASRDVSVYNTNMGLKFRQLMFHI 121
Qy 138 SCLTFGRVETIYLVSGFWIRTPPAYRPNAPILSTLPTTIVRRGRSPRRTPSPRR 197
Db 122 SCLTFGRVETIYLVSGFWIRTPPAYRPNAPILSTLPTTIVRRGRSPRRTPSPRR 181
Qy 198 RRSQSPRRRSQSPRR 215
Db 182 RRSQSPRRRSQSPRR 199

RESULT 4

US-08-968-747-2
Sequence 2, Application US/08968747
Patent No. 6060595
GENERAL INFORMATION:
APPLICANT: Scaglioni et al.
TITLE OF INVENTION: INHIBITION OF VIRAL REPLICATION
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA

COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/968,747
FILING DATE: 03-SEP-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 08472/705001
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 193 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-968-747-2

Query Match 87.2%; Score 1000.5; DB 3; Length 193;
Best Local Similarity 96.0%; Pred. No. 4.7e-102; Indels 5; Gaps 1;
Matches 190; Conservative 1; Mismatches 2;

QY 18 SKLCIGWMDIDPYKRGATVELLSFLPSDFPSVVDLDTASALYREALSEPHCS 77
DB 1 SKLCIGWMDIDPYKRGATVELLSFLPSDFPSVVDLDTASALYREALSEPHCS 60
QY 78 HTALROAILCGELMTLATWGVNLEDPFRGDASRDLYSVYNTMGLKROLIMFHI 137
DB 61 HTALROAILCGELMTLATWGVNLEDP-----ASRDLYSVYNTMGLKROLIMFHI 115
QY 138 SCLFGRFETVLEYVSGVWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRTPSPDR 197
DB 116 SCLFGRFETVLEYVSGVWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRTPSPDR 175
QY 198 RRSQSPRRRSQSPRRPQC 215
DB 176 RRSQSPRRRSQSPRRPQC 193

RESULT 5
US-08-105-483-217
Sequence 217, Application US/08105483
Patent No. 5494807
GENERAL INFORMATION:
APPLICANT: Paoletti, Enzo
TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE
NUMBER OF SEQUENCES: 462
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford
ADDRESSEE: c/o William S. Frommer
STREET: 530 Fifth Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/105,483

FILING DATE: 12-AUG-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/847,951
FILING DATE: 06-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2400
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 217:
SEQUENCE CHARACTERISTICS:
LENGTH: 346 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-105-483-217

Query Match 84.7%; Score 972; DB 1; Length 346;
Best Local Similarity 86.7%; Pred. No. 1.5e-98;
Matches 189; Conservative 2; Mismatches 6; Indels 16; Gaps 2;

QY 3 LSSIFRIGDPTVOASKCLGWLGMMDIDPYKRGATVELLSFLPSDFPSVVDLDTAS 62
DB 150 LSSIFRIGDPTVOASKCLGWLGMMDIDPYKRGATVELLSFLPSDFPSVVDLDTAS 198
QY 63 ALYRBALSEPHCSPPHTALROAILCGELMTLATWGVNLEDPFRGDASRDLYSVYN 122
DB 199 ALYRBALSEPHCSPPHTALROAILCGELMTLATWGVNLEDP-----ASRDLYSVYN 253
QY 123 TMGLKROLIMFHSCLTGRFETVLEYVSGVWIRTPPAYRPPNAPILSTLPTTVVR 182
DB 254 TMGLKROLIMFHSCLTGRFETVLEYVSGVWIRTPPAYRPPNAPILSTLPTTVVR 313
QY 183 RRGSPRRRTPSPRRRSQSPRRRSQSPRRPQC 215
DB 314 RRGSPRRRTPSPRRRSQSPRRRSQSPRRPQC 346

RESULT 6
US-08-709-209-217
Sequence 217, Application US/08709209
Patent No. 5762938
GENERAL INFORMATION:
APPLICANT: Paoletti, Enzo
TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE
NUMBER OF SEQUENCES: 462
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford
ADDRESSEE: c/o William S. Frommer
STREET: 530 Fifth Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/709,209
FILING DATE: 21-AUG-1996
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/105,483
FILING DATE: 12-AUG-1993
APPLICATION NUMBER: US 07/847,951
FILING DATE: 06-MAR-1992
ATTORNEY/AGENT INFORMATION:

```

NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 44310-2400
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 217:
SEQUENCE CHARACTERISTICS:
LENGTH: 346 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-709-209-217

Query Match 84.7% Score 972; DB 1; Length 346;
Beet Local Similarity 88.7% Pred.No.1.5e-98;
Matches 189; Conservative 2; Mismatches 6; Indels 16; Gaps 2

OY 3 LSSISRIQDPVQVQAKCLGWLGMGMDIDPYKEFGATVELLSFLPSDFPPSYVDLDTAS 62
| | | | |
Db 150 ISSIARIQDDPYT-----NMDDIPYKEFGATVELLSFLPSDFPPSYVDLDTAS 198
| | | | |

OY 63 ALYREALSEPHCSPHHTALROAILCWSGIMTLATWGVNLDDPEFRGDSRDLVVSYN 122
| | | | |
Db 199 ALYREALSEPHCSPHHTALROAILCWSGIMTLATWGVNLDDP-----ASRDLVVSYN 253
| | | | |

OY 123 TMMGLKROLNMFHSICLTPGKRTVIEYLVRSQWARTPTAPYPPNAPILSTLPTTVR 182
| | | | |
Db 254 TMMGLKROLNMFHSICLTPGKRTVIEYLVRSQWARTPTAPYPPNAPILSTLPTTVR 313
| | | | |

OY 183 RRGSRPRRTPSPRRRRSOPRRRRSOPRRSOPC 215
| | | | |
Db 314 RRGSRPRRTPSPRRRRSOPRRRRSOPRRRRSOPC 346
| | | | |

RESULT 7
US-08-458-101-217
Sequence 217, Application US/08458101
Patent No. 5765399
GENERAL INFORMATION:
APPLICANT: Paoletti, Enzo
APPLICANT: Perkus, Marion E.
APPLICANT: Taylor, Jill
APPLICANT: Tartaglia, James
APPLICANT: No. 5765399con, Elizabeth K.
APPLICANT: Riviere, Michel
APPLICANT: de Taisne, Charles
APPLICANT: Limbach, Keith J.
APPLICANT: Johnson, Gerard P.
APPLICANT: Pincus, Steven E.
APPLICANT: Cox, William I.
APPLICANT: Audonnet, Jean-Christophe Francis
APPLICANT: Gettigs, Russell Robert
TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE
TITLE OF INVENTION: STRAIN
NUMBER OF SEQUENCES: 467
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford
ADDRESSEE: c/o William S. Frommer
STREET: 530 Fifth Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,101
FILING DATE: 01-JUN-1995
CLASSIFICATION: 424

```

```

ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2740
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 217:
SEQUENCE CHARACTERISTICS:
LENGTH: 346 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-458-101-217

Query Match      84.7%  Score 972;  DB 1;  Length 346;
Blast Local Similarity 88.7%  Pred. No. 1.5e-96;
Matches 189;  Conservative 2;  Mismatches 6;  Indels 16;  Gaps 2.

3  LSSISRIQDPVQAASKLCLGWLGMQMDIDPYKEFGATVELLSFLPSDFPSPYRDDLLDTAS 62
150 ISSIARIGDPT-----NMDDIPYKRFATYELLSFLPSDFPSPYRDDLLDTAS 198

63 ALYREALESPKCHSPHTALRQAILCWELMNTLATWGVNLDEPEFRGDASGLDYVSYN 122
199 ALYREALESPKCHSPHTALRQAILCWELMNTLATWGVNLDEP-----ASRDLYVSYN 253

123 TMMGLKFRQLMFHISCLTFGREYIYEVYVSGWIRTPPAVRPNAPILSTLPETTVR 182
254 TMMGLKFRQLMFHISCLTFGREYIYEVYVSGWIRTPPAVRPNAPILSTLPETTVR 313

183 RRGSPRRRTSPRRRRSOSPPRRRSOSRPPQC 215
314 RRGSPRRRTSPRRRRSOSPPRRRSOSRPPQC 346

```

```

RESULT 8
5196194-13
; Patent No. 5196194
; APPLICANT: RUTTER, WILLIAM J.; GOODMAN, HOWARD M.
; TITLE OF INVENTION: VACCINES CONTAINING HEPATITIS B S-PROTEIN
; NUMBER OF SEQUENCES: 21
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/679,621
; FILING DATE: 7-DEC-1984
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 513,055
; FILING DATE: 12-JUL-1983
; APPLICATION NUMBER: 107,267
; FILING DATE: 21-DEC-1979
; APPLICATION NUMBER: 41,909
; FILING DATE: 24-MAY-1979
; SEQ ID NO:13:
; LENGTH: 211
5196194-13

Query Match      83.1%; Score 954.5; DB 6; Length 211;
Best Local Similarity 91.2%; Pred. No. 66-97; Indels 9; Gaps
Matches 187; Conservative 3; Mismatches 6

QY 13 PTQASALCLGIMGMDIDPYKEFGATVELSLPSDFPFSVYRDLDITASALYREALLESP 72
14 PTQASRLCLGIMGMDIDPYKEFGATVELSLPSDFPFSVYRDLDITASAGYREALLESP 72
73 EHCSPHHTALROAILCWMGELMTLATVGNVLEDDPEFRGASRDLDVSYVNTNMGKFRQL 132
73 EHCSPHHTALROAILCWMGELMTLATVGNVLEDDPEFRGASRDLDVSYVNTNMGKFRQL 126
QY 133 LWHHISLTLTGRTVLEVLVSFGWVIRTPAPVPPNAPILSTLPTETTVYRR--RGRSPRR 190
127 LWHHISLTLTGRTVLEVLVSFGWVIRTPAPVPPNAPILSTLPTETTVYRR--RGRSPRR 186
QY 191 RTPSPRRRRSGSPRRRRSGRFPQC 215

```

Db 167 RTSPRRRRSPRRRRSOSRESQC 211

RESULT 9

US-09-248-588-2

Sequence 2. Application US/09248588

Patent No. 6231864

GENERAL INFORMATION:

APPLICANT: Birkett, Ashley J.

TITLE OF INVENTION: Strategically Modified Hepatitis B Core Proteins and

FILE REFERENCE: SYN-101 4564/59529

CURRENT APPLICATION NUMBER: US/09/248,588

EARLIER FILING DATE: 1999-02-11

EARLIER APPLICATION NUMBER: 60/074537

NUMBER OF SEQ ID NOS: 113

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO: 2

LENGTH: 183

TYPE: PRT

ORGANISM: Hepatitis B virus

US-09-248-588-2

Query Match

Best Local Similarity 96.8%; Score 951.5; DB 4; Length 183;

Matches 182; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

Db 1

Qy 28 MDIDPKKFGATVETLSFLPSDFPVSVDLIDTASALYREALSPHCSPHHTALRQAIL 87

Db 1 MDIDPKKFGATVETLSFLPSDFPVSVDLIDTASALYREALSPHCSPHHTALRQAIL 60

Qy 88 CMGELMTLATWGVNLEDPFRGDSRDIVSVYNTNMGLKFRQLMFHISCLTFGRGV 147

Db 61 CMGELMTLATWGVNLEDPFRGDSRDIVSVYNTNMGLKFRQLMFHISCLTFGRGV 115

Qy 148 IEYVSPGVWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTPSPRRRSOSPPRRR 207

Db 116 IEYVSPGVWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTPSPRRRSOSPPRRR 175

Qy 208 SOSRESQC 215

Db 176 SOSRESQC 183

RESULT 10

PCT-US96-10602-12

Sequence 12. Application PC/TUS9610602

GENERAL INFORMATION:

APPLICANT: The General Hospital Corporation

TITLE OF INVENTION: INHIBITION OF HEPATITIS B REPLICATION

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US96/10602

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/017,814

FILING DATE: 20-JUN-1995

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Clark, Paul T.

REGISTRATION NUMBER: 30,162

REFERENCE/DOCKET NUMBER: 00786/282001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617/542-5070

TELEFAX: 617/542-8906

TELEX: 200154

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 183 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

PCT-US96-10602-12

Query Match

Best Local Similarity 81.8%; Score 938.5; DB 5; Length 183;

Matches 180; Conservative 0; Mismatches 3; Indels 5; Gaps 1;

Db 1

Qy 28 MDIDPKKFGATVETLSFLPSDFPVSVDLIDTASALYREALSPHCSPHHTALRQAIL 87

Db 1 MDIDPKKFGATVETLSFLPSDFPVSVDLIDTASALYREALSPHCSPHHTALRQAIL 60

Qy 88 CMGELMTLATWGVNLEDPFRGDSRDIVSVYNTNMGLKFRQLMFHISCLTFGRGV 147

Db 61 CMGELMTLATWGVNLEDPFRGDSRDIVSVYNTNMGLKFRQLMFHISCLTFGRGV 115

Qy 148 IEYVSPGVWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTPSPRRRSOSPPRRR 207

Db 116 IEYVSPGVWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTPSPRRRSOSPPRRR 175

Qy 208 SOSRESQC 215

Db 176 SOSRESQC 183

RESULT 11

US-08-968-747-20

Sequence 20. Application US/08968747

Patent No. 6060595

GENERAL INFORMATION:

APPLICANT: Scaglioni et al.

TITLE OF INVENTION: INHIBITION OF VIRAL REPLICATION

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: Windows

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/968,747

FILING DATE: 03-SEP-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Frazer, Janis K.

REGISTRATION NUMBER: 34,819

REFERENCE/DOCKET NUMBER: 08472/705001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617/542-5070

TELEFAX: 617/542-8906

TELEX: 200154

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 183 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-968-747-20

Query Match 81.6%; Score 936.5; DB 3; Length 183;
 Best Local Similarity 95.7%; Pred. No. 4,7e-95;
 Matches 180; Conservative 1; Mismatches 2; Indels 5; Gaps 1;

QY 28 MDIDPYKEFGATVLLSLFSDPFPSVRDLDTASALVREALSPHCSPHHTALROAIL 87
 DB 1 MDIDPYKEFGATVLLSLFSDPFPSVRDLDTASALVREALSPHCSPHHTALROAIL 60
 QY 88 CWGELMTLATWGVNLEDEFRGDASRDVLVSYVNTNMGKFKROLMLFHSCLTFGRRTV 147
 DB 61 CWGELMTLATWGVNLEDEFRGDASRDVLVSYVNTNMGKFKROLMLFHSCLTFGRRTV 115
 QY 148 IEVLVFGWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRR 207
 DB 116 IEVLVFGWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRR 175
 QY 208 SOSRPOC 215
 DB 176 SOSRESOC 183

RESULT 12

PCT-US96-10602-6
 Sequence 6, Application PC/TUS9610602

GENERAL INFORMATION:
 APPLICANT: The General Hospital Corporation
 TITLE OF INVENTION: INHIBITION OF HEPATITIS B REPLICATION
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson P.C.
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02110-2804
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US96/10602
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/017,814
 FILING DATE: 20-JUN-1995
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Clark, Paul T.
 REGISTRATION NUMBER: 30,162
 REFERENCE/DOCKET NUMBER: 00786/282001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617/542-5070
 TELEFAX: 617/542-8906
 TELEX: 200154
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 397 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 PCT-US96-10602-6

Query Match 81.3%; Score 933.5; DB 5; Length 397;
 Best Local Similarity 97.3%; Pred. No. 3e-94;
 Matches 179; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 28 MDIDPYKEFGATVLLSLFSDPFPSVRDLDTASALVREALSPHCSPHHTALROAIL 87
 DB 1 MDIDPYKEFGATVLLSLFSDPFPSVRDLDTASALVREALSPHCSPHHTALROAIL 60
 QY 88 CWGELMTLATWGVNLEDEFRGDASRDVLVSYVNTNMGKFKROLMLFHSCLTFGRRTV 147
 DB 61 CWGELMTLATWGVNLEDEFRGDASRDVLVSYVNTNMGKFKROLMLFHSCLTFGRRTV 115
 QY 148 IEVLVFGWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRR 207
 DB 116 IEVLVFGWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRR 175
 QY 208 SOSR 211
 DB 176 SOSR 179

RESULT 13

US-07-739-642-2
 Sequence 2, Application US/07739642

Patent No. 5173427
 GENERAL INFORMATION:
 APPLICANT: Mallonee, Richard L.
 TITLE OF INVENTION: Vectors and Hosts with Increased
 NUMBER OF SEQUENCES: 24
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Richard R. Rodrick
 STREET: 1 Beeton Drive
 CITY: Franklin Lakes
 STATE: New Jersey
 COUNTRY: U.S.A.
 ZIP: 07417-1880
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/739,642
 FILING DATE: 19910801
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Stierwalt, Brian K.
 REGISTRATION NUMBER: 33,213
 REFERENCE/DOCKET NUMBER: P-2272
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 201-848-5317
 TELEFAX: 201-848-9228
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 185 amino acids
 TYPE: AMINO ACID
 TOPOLOGY: linear
 US-07-739-642-2

Query Match 80.4%; Score 922.5; DB 1; Length 185;
 Best Local Similarity 93.7%; Pred. No. 1.6e-93;
 Matches 178; Conservative 2; Mismatches 3; Indels 7; Gaps 2;

QY 28 MDIDPYKEFGATVLLSLFSDPFPSVRDLDTASALVREALSPHCSPHHTALROAIL 87
 DB 1 MDIDPYKEFGATVLLSLFSDPFPSVRDLDTASALVREALSPHCSPHHTALROAIL 60
 QY 88 CWGELMTLATWGVNLEDEFRGDASRDVLVSYVNTNMGKFKROLMLFHSCLTFGRRTV 147
 DB 61 CWGELMTLATWGVNLEDEFRGDASRDVLVSYVNTNMGKFKROLMLFHSCLTFGRRTV 115
 QY 148 IEVLVFGWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRR 205
 DB 116 IEVLVFGWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRR 175
 QY 206 RSOSRPOC 215

Db 176 RRSQRESQC 185

RESULT 14

US-07-739-643-2
Sequence 2, Application US/07739643

PATENT No. 5175094

GENERAL INFORMATION:

APPLICANT: Mallonee, Richard L.

TITLE OF INVENTION: Increased Expression of HBcAg

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESSER: Richard R. Rodrick

STREET: 1 Becton Drive

CITY: Franklin Lakes

STATE: New Jersey

COUNTRY: U.S.A.

ZIP: 07417-1880

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/739,643

FILING DATE: 19910801

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Stierwalt, Brian K.

REGISTRATION NUMBER: 33,213

REFERENCE/DOCKET NUMBER: P-2090

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-848-5317

TELEFAX: 201-848-9228

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 185 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

US-07-739-643-2

Query Match

Best Local Similarity 93.7%; Score 922.5; DB 1; Length 185;

Matches 176; Conservative 2; Mismatches 3; Indels 7; Gaps 2;

Db 28 MDIDYKRGATVVELLSFLPSDFPSVVDLDTASALYREALSPHSCSPHTALRQAIL 87

1 MDIDYKRGATVVELLSFLPSDFPSVVDLDTASALYREALSPHSCSPHTALRQAIL 60

Qy 88 CWGELMTLATVGVNLEDPERFGDASRDLYVSYVNTNMGKLFQQLMFIHISCLTFGRRTV 147

61 CWGELMTLATVGVNLEDPERFGDASRDLYVSYVNTNMGKLFQQLMFIHISCLTFGRRTV 115

Qy 148 IEYVSGVWIRTPPAYRPPNAPILSTLPETTVVR--RGRSPRRTPSPRRRSQSPRR 205

Db 116 IEYVSGVWIRTPPAYRPPNAPILSTLPETTVVR--RGRSPRRTPSPRRRSQSPRR 175

Qy 206 RRSQRESQC 215

Db 176 RRSQRESQC 185

RESULT 15

US-07-739-142-2

Sequence 2, Application US/07739142

PATENT No. 5175272

GENERAL INFORMATION:

APPLICANT: Mallonee, Richard L.

TITLE OF INVENTION: DNA Sequences with Increased Expression

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESSER: Richard R. Rodrick

STREET: 1 Becton Drive

CITY: Franklin Lakes

STATE: New Jersey

COUNTRY: U.S.A.

ZIP: 07417-1880

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/739,142

FILING DATE: 19910801

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Stierwalt, Brian K.

REGISTRATION NUMBER: 33,213

REFERENCE/DOCKET NUMBER: P-2271

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-848-5317

TELEFAX: 201-848-9228

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 185 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

US-07-739-142-2

Query Match

Best Local Similarity 93.7%; Score 922.5; DB 1; Length 185;

Matches 176; Conservative 2; Mismatches 3; Indels 7; Gaps 2;

Db 28 MDIDYKRGATVVELLSFLPSDFPSVVDLDTASALYREALSPHSCSPHTALRQAIL 87

1 MDIDYKRGATVVELLSFLPSDFPSVVDLDTASALYREALSPHSCSPHTALRQAIL 60

Qy 88 CWGELMTLATVGVNLEDPERFGDASRDLYVSYVNTNMGKLFQQLMFIHISCLTFGRRTV 147

61 CWGELMTLATVGVNLEDPERFGDASRDLYVSYVNTNMGKLFQQLMFIHISCLTFGRRTV 115

Qy 148 IEYVSGVWIRTPPAYRPPNAPILSTLPETTVVR--RGRSPRRTPSPRRRSQSPRR 205

Db 116 IEYVSGVWIRTPPAYRPPNAPILSTLPETTVVR--RGRSPRRTPSPRRRSQSPRR 175

Qy 206 RRSQRESQC 215

Db 176 RRSQRESQC 185

Search completed: February 3, 2003, 09:27.10

Job time: 14.1098 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 09:15:41 ; Search time 34.4599 Seconds
(without alignments)
831,368 Million cell updates/sec

Title: US-09-890-752A-2

Perfect score: 1148
Sequence: 1 MFLSSIFSRKIGDPTVQASKL.....RRRSQSPRRRSQSRPQC 215

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: A_Geneseq.101002:*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1148	100.0	215	21	AA10597
2	1048.5	91.3	212	23	AA080921
3	1040.5	90.6	212	19	AA050250
4	1040.5	90.6	212	23	AA019793
5	1040.5	90.6	212	23	AA019898
6	1040.5	90.6	212	23	AA019708
7	1037.5	90.4	212	22	AA066974
8	1037.5	90.4	212	23	AA080914
9	1037.5	90.4	212	23	AA080915
10	1037.5	90.4	212	23	AA080916

11	1035.5	90.2	212	23	AA080919	Hepatitis B virus
12	1035.5	90.2	212	23	AA080920	Hepatitis B virus
13	1034.5	90.1	212	23	AA080922	Hepatitis B virus
14	1034.5	90.1	212	23	AA080923	Hepatitis B virus
15	1034.5	90.1	212	23	AA080924	Hepatitis B virus
16	1031.5	89.9	212	23	AA080928	Hepatitis B virus
17	1031.5	89.9	212	23	AA080929	Hepatitis B virus
18	1030.5	89.8	212	23	AA080930	Hepatitis B virus
19	1030.5	89.8	212	23	AA080931	Hepatitis B virus
20	1030.5	89.8	212	23	AA080932	Hepatitis B virus
21	1029.5	89.7	212	23	AA080933	Hepatitis B virus
22	1029.5	89.7	212	23	AA080934	Hepatitis B virus
23	1028.5	89.6	212	23	AA080937	Hepatitis B virus
24	1028.5	89.6	212	23	AA080938	Hepatitis B virus
25	1028.5	89.6	212	23	AA080939	Hepatitis B virus
26	1023.5	89.2	212	23	AA080940	Hepatitis B virus
27	1022.5	89.1	212	23	AA080941	Hepatitis B virus
28	1020.5	88.9	212	23	AA080901	Hepatitis B virus
29	1019.5	88.8	212	22	AA066975	Hepatitis B virus
30	1019.5	88.8	212	23	AA080936	Hepatitis B virus
31	1015.5	88.5	194	19	AA050242	Hepatitis B virus
32	1015.5	88.5	212	23	AA080942	Hepatitis B virus
33	1014.5	88.4	212	23	AA080902	Hepatitis B virus
34	1011.5	88.1	212	23	AA080907	Hepatitis B virus
35	1011.5	88.1	214	10	AA090702	Hepatitis B virus
36	1008.5	87.8	212	23	AA080908	Deduced amino acid
37	1007.5	87.8	212	22	AA066972	Hepatitis B virus
38	1007.5	87.8	212	22	AA066973	Hepatitis B virus
39	1007.5	87.8	212	22	AA066974	Hepatitis B virus
40	1002.5	87.3	214	23	AA080931	Hepatitis B virus
41	1000.5	87.2	193	19	AA050241	Hepatitis B virus
42	997.5	87.1	198	19	AA050252	Human hepatitis B
43	997.5	86.9	212	21	AA044350	Hepatitis B virus
44	996.5	86.8	212	23	AA080918	Hepatitis B virus
45	996.5	86.8	214	9	AA080961	HBV core antigen e

ALIGNMENTS

RESULT 1
AA010597 standard; Protein; 215 AA.
ID AA010597
XX
AC AA010597;
XX
DT 08-JAN-2001 (first entry)
XX
DE HBV fusion protein comprising HBcAg and RGD.
XX
KW Fusion protein; protein coat; virus-specific packaging signal; pol;
KW virus protein; cell permeability; cell-specific binding site; LHB;
KW large surface protein; core antigen; gene therapy.
XX
OS Hepatitis B virus.
OS Synthetic.
PN MO200046376-A2.
XX
PD 10-AUG-2000.
XX
PF 04-FEB-2000; 2000WO-DE00363.
XX
PR 05-FEB-1999; 99DE-1004800.
XX
PA (HILD/) HILD E.
XX
PI Hildt E, Hofschneider P;
XX
XX WPI, 2000-514959/46.
XX N-PSDB; AA017135.
XX
PT Particle for cell-specific gene delivery, useful in gene therapy,

AC AAM50250;
 XX 28-SEP-1998 (first entry)
 XX Hepatitis B virus precursor p25 polypeptide.
 DE Hepatitis B virus precursor p25 polypeptide.
 XX Viral replication; inhibitor; HBV, nucleocapsid; gene therapy;
 KM hepatocyte; liver; p25 protein.
 XX Hepatitis B virus.
 OS
 PN WO9809649-A1.
 XX 12-MAR-1998.
 XX
 PD 03-SEP-1997; 97WO-US15500.
 XX
 XX 03-SEP-1996; 96US-0025370.
 XX
 XX (GENO) GEN HOSPITAL CORP.
 PA
 PI Melegari M, Scaglioni PP, Wands JR;
 DR WPI; 1998-193325/17.
 XX
 XX DNA encoding proteins which can be incorporated with wild type
 PT nucleocapsid subunit(s) into a viral nucleocapsid - useful for
 PT inhibition of viral replication, especially hepatitis B virus
 XX
 PS Claim 15; Page 35; 60pp; English.

This polypeptide comprises the hepatitis B virus (HBV) 25 kDa (p25)
 CC protein that is encoded by the full-length HBV precore gene.
 CC Evidence is provided that HBV replication is inhibited in the
 CC presence of high levels of HBV precore or precore-related proteins.
 CC These proteins can be incorporated into HBV nucleocapsids along
 CC with the p21 core protein (see AAM50251), which is the usual
 CC nucleocapsid component, and thereby render the nucleocapsids
 CC deficient in encapsidating HBV pregenomic RNA. Thus, over-
 CC expression of the precore proteins, or certain variants of them,
 CC leads to transdominant inhibition of HBV replication. Suitable
 CC inhibitory proteins include p25, p22 (see AAM50241), Met-p22 (see
 CC AAM50242), p18 (see AAM50236), Met-p18 (see AAM50237) and Met-p18-Het
 CC (see AAM50238). The inhibitory proteins can be produced by
 CC recombinant methods using claimed expression vectors and host
 CC cells. They can be provided exogenously to the target cells for
 CC use in inhibiting HBV replication. Alternatively, a nucleic acid
 CC construct that directs overexpression of an inhibitory protein in
 CC target cells is used for the gene therapy of HBV infection.

Sequence 212 AA;
 SQ

Query Match 90.6%; Score 1040.5; DB 19; Length 212;
 Best Local Similarity 97.0%; Pred. No. 2.5e-98;
 Matches 197; Conservative 0; Mismatches 1; Indels 5; Gaps 1;
 XX
 QY 13 PTVQASKLCLGMLMGMDIDPYKFGATVLLSFLPSDFPSVRDLIDTASALYREALBSP 72
 DB 15 PTVQASKLCLGMLMGMDIDPYKFGATVLLSFLPSDFPSVRDLIDTASALYREALBSP 74
 QY 73 EHCSPHHTALROAILCWMGLMTLATWGVNLEDPFRGDASRDVLVSVYNTMGLKFRQL 132
 DB 75 EHCSPHHTALROAILCWMGLMTLATWGVNLEDPFRGDASRDVLVSVYNTMGLKFRQL 129
 QY 133 LMFHISCLTFGRRETVIEVLVSGVWIRTPPAYRPNPAILSTLPTTVVRRGRSPRRRT 192
 DB 130 LMFHISCLTFGRRETVIEVLVSGVWIRTPPAYRPNPAILSTLPTTVVRRGRSPRRRT 189
 QY 193 PSPRRRSQSPRRRSQSPREPOC 215
 DB 190 PSPRRRSQSPRRRSQSPREPOC 212

RESULT 4
 AAE19793
 ID AAE19793 standard; Protein, 212 AA.
 XX
 XX AAE19793;
 AC
 XX 18-JUN-2002 (first entry)
 DT
 DE Hepatitis B virus core antigenic protein (HBcAg).
 XX
 XX Hepatitis B virus core antigen; HBcAg; prophylactic; viral hepatitis;
 KM therapeutic; vaccine; acquired immune deficiency syndrome; influenza;
 KM polio; herpes; rabies; AIDS; foot-and-mouth disease.
 XX
 OS Hepatitis B virus.
 PN WO200177158-A1.
 XX
 XX 18-OCT-2001.
 XX
 XX 09-APR-2001; 2001WO-GB01607.
 XX
 XX 07-APR-2000; 2000EP-0107118.
 XX
 XX (MEDVA) MEDEVA EURO LTD.
 XX
 XX Gehin A, Gilbert R, Stuart D, Rowlands D;
 DR WPI; 2002-239995/29.
 DR N-PSDB; AAD31509.
 XX
 XX Hepatitis B (HB) core antigen fusion proteins, useful as vaccines for
 PT the prophylactic or therapeutic treatment of humans or animals against
 PT e.g. HB virus, viral hepatitis, hepatitis C virus, influenza, or
 PT foot-and-mouth disease -
 XX
 XX Disclosure; Page 24-25; 27pp; English.

The present invention relates to hepatitis B virus (HBV) core antigen
 CC (HBcAg) fusion proteins and polynucleotides encoding such proteins.
 CC Sequences of the invention are useful in methods of prophylactic or
 CC therapeutic vaccination or to manufacture medicaments for prophylactic
 CC or therapeutic vaccination of the human or animal body against HBV,
 CC e.g. against viral hepatitis. They are also useful as a prophylactic
 CC vaccine against e.g. hepatitis C virus, influenza, polio, herpes,
 CC rabies, acquired immune deficiency syndrome (AIDS) or foot-and-mouth
 CC disease. The present sequence is hepatitis B virus core antigenic
 CC protein (HBcAg).
 XX
 XX

Sequence 212 AA;
 SQ

Query Match 90.6%; Score 1040.5; DB 23; Length 212;
 Best Local Similarity 97.0%; Pred. No. 2.5e-98;
 Matches 197; Conservative 0; Mismatches 1; Indels 5; Gaps 1;
 XX
 QY 13 PTVQASKLCLGMLMGMDIDPYKFGATVLLSFLPSDFPSVRDLIDTASALYREALBSP 72
 DB 15 PTVQASKLCLGMLMGMDIDPYKFGATVLLSFLPSDFPSVRDLIDTASALYREALBSP 74
 QY 73 EHCSPHHTALROAILCWMGLMTLATWGVNLEDPFRGDASRDVLVSVYNTMGLKFRQL 132
 DB 75 EHCSPHHTALROAILCWMGLMTLATWGVNLEDPFRGDASRDVLVSVYNTMGLKFRQL 129
 QY 133 LMFHISCLTFGRRETVIEVLVSGVWIRTPPAYRPNPAILSTLPTTVVRRGRSPRRRT 192
 DB 130 LMFHISCLTFGRRETVIEVLVSGVWIRTPPAYRPNPAILSTLPTTVVRRGRSPRRRT 189

Qy 193 PSPRRRSQSPRRRSQSPREPQC 215
 |||||
 Db 190 PSPRRRSQSPRRRSQSPRESQC 212

RESULT 5

AAE19898
 ID AAE19898 standard; Protein; 212 AA.

XX AAE19898;

XX 18-JUN-2002 (first entry)

XX Hepatitis B virus C antigen (HBcAg) and E antigen (HEAg) sequence.

XX Hepatitis B virus; HBV; infection; virucide; fungicide; antibacterial;

XX cytoskeletal; immunostimulant; vaccine; ribavirin; immune response; cancer.

XX Hepatitis B virus.

XX WO200213855-A2.

XX 21-FEB-2002.

XX 15-AUG-2001; 2001WO-1B01808.

XX 17-AUG-2000; 2000US-225767P.

XX 29-AUG-2000; 2000US-229175P.

XX 03-NOV-2000; 2000US-0705547.

XX (TRIP-) TRIPEP AB.

XX Salberg M, Hultgren C;

XX WPI; 2002-241837/29.

XX Vaccine compositions for treating and preventing disease, preferably

XX hepatitis C virus infection, comprises ribavirin and antigen that has

XX epitope present in hepatitis C virus -

XX Claim 11; Page 82; 120pp; English.

XX The invention relates to a composition comprising ribavirin and an

XX antigen preferably non structural 3 protein (NS3)/4A fragment of

XX hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV

XX sequence. The composition is useful for enhancing an immune response to

XX a hepatitis C antigen in humans, domestic, sport or pet species and as

XX vaccines for treating and preventing HCV infections. The composition is

XX also useful for treating viral, bacterial, fungal diseases and cancer.

XX The present sequence is Hepatitis B virus C antigen (HBcAg) and E antigen

XX (HEAg) sequence.

XX Sequence 212 AA;

XX Query Match 90.6%; Score 1040.5; DB 23; Length 212;

XX Best Local Similarity 97.0%; Pred. No. 2.5e-98; Indels 5; Gaps 1;

XX Matches 197; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

Qy 13 PTVQASKLCLGMLGMDIDPYKEFGATVLLSFLPSDFPSVRDLDTASALYREALESP 72

Db 15 PTVQASKLCLGMLGMDIDPYKEFGATVLLSFLPSDFPSVRDLDTASALYREALESP 74

RESULT 6

AAE17018
 ID AAE17018 standard; Protein; 212 AA.

XX AAE17018;

XX 18-APR-2002 (first entry)

XX Hepatitis B virus (HBV) core antigen (HBcAg).

XX Hepatitis B virus; HBV; core antigen; HBcAg; immune system; typhoid;

XX prophylactic; gene therapy; vaccine; hepatitis A virus; HAV; herpes;

XX hepatitis C virus; HCV; influenza; foot-and-mouth disease; diarrhoea;

XX tuberculosis; polio; rabies; acquired immunodeficiency syndrome; AIDS;

XX dengue fever; yellow fever; malaria; whooping cough; salmonellosis;

XX food poisoning; meningitis; gonorrhea; antiviral; antibacterial;

XX anti-protozoal.

XX Hepatitis B virus.

XX WO200198333-A2.

XX 27-DEC-2001.

XX 22-JUN-2001; 2001WO-GB02817.

XX 22-JUN-2000; 2000GB-0015308.

XX 06-OCT-2000; 2000GB-0024544.

XX (CELL-) CELTECH PHARM LTD.

XX Page M, Li J, Pumpens P;

XX WPI; 2002-098223/13.

XX N-PSDB; AAD27422.

XX New proteins comprising a modified hepatitis B core antigen, useful as

XX a vaccine in prophylactic or therapeutic vaccination of the human or

XX animal body, particularly against hepatitis B virus infection -

XX Disclosure; Page 39-40; 40pp; English.

XX The invention relates to modified proteins comprising hepatitis B virus

XX (HBV) core antigen (HBcAg) wherein one or more of the four arginine

XX repeats has been deleted and the protein comprising the C-terminal

XX cysteine of HBcAg. The deleted region may be replaced by an epitope

XX from a protein other than HBcAg, in which case the HBcAg acts as a

XX carrier to present the epitope to the immune system. This chimeric

XX protein or its nucleic acid is useful as a vaccine or in a method of

XX prophylactic or therapeutic vaccination of the human or animal body,

XX particularly against HBV. The nucleic acid encoding the protein may

XX be used in gene therapy or DNA vaccination protocols. The chimeric

XX protein or its nucleic acid may also be used as the basis of a

XX prophylactic vaccine against a range of diseases, e.g. HBV, hepatitis

XX A virus (HAV), hepatitis C virus (HCV), influenza, foot-and-mouth

XX disease, polio, herpes, acquired immunodeficiency syndrome, AIDS,

XX (AIDS), dengue fever, yellow fever, malaria, tuberculosis, whooping

XX cough, salmonellosis, typhoid, food poisoning, diarrhoea, meningitis

XX or gonorrhea. The present sequence is Hepatitis B virus core antigen

XX (HBcAg).

XX Sequence 212 AA;

XX Query Match 90.6%; Score 1040.5; DB 23; Length 212;

XX Best Local Similarity 97.0%; Pred. No. 2.5e-98; Indels 5; Gaps 1;

XX Matches 197; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

Qy 13 PTVQASKLCLGMLGMDIDPYKEFGATVLLSFLPSDFPSVRDLDTASALYREALESP 72

Db 15 PTVQASKLCLGMLGMDIDPYKEFGATVLLSFLPSDFPSVRDLDTASALYREALESP 74

QY 73 EHCSPHHTALROAILCQGLMTLTA TWGVNLEDPFRGASDLVSVYNTNMGLKFRQL 132
 DB 75 EHCSPHHTALROAILCQGLMTLTA TWGVNLEDP-----ASRDLVSVYNTNMGLKFRQL 129
 QY 133 LMFHISCLTFGRRETVIEYLVSPGVWIRTPPAYRPNAPILSTLPETTVYRRGRSPRRRT 192
 DB 130 LMFHISCLTFGRRETVIEYLVSPGVWIRTPPAYRPNAPILSTLPETTVYRRGRSPRRRT 189
 QY 193 PSPRRRSQSPRRRSQSPRRSQ 215
 DB 190 PSPRRRSQSPRRRSQSPRRSQ 212

RESULT 7

AA66924
 ID AAG66924 standard; Protein; 212 AA.

AC AAG66924;

DT 19-OCT-2001 (first entry)

DE HBV genotype D preCore and Core polypeptide.

KW Hepatitis B virus; HBV; preCore; Core; preS1; preS2; HBS; HBs; HBpol;
 KM HBsAg; antiviral; vaccine; genotype G; genotyping; HBcAg; HBeAg.

OS Hepatitis B virus.

FN WO200140279-A2.

PD 07-JUN-2001.

PF 20-NOV-2000; 2000WO-EP11526.

PR 03-DEC-1999; 99EP-0870252.
 PR 07-DEC-1999; 99US-0169287.

PA (INNO-) INNOGENETICS NV.

PI Stuyver L, Van Geyt C, De Gendt S;

DR WPI; 2001-374785/39.

PT Novel isolated and/or purified hepatitis B virus polypeptide and
 PT polynucleotide sequences that are phylogenetically different from HBV
 PT genotype A-F molecules, useful for HBV diagnosis, prophylaxis and
 PT therapy -

PS Example 3; Fig 5; 94pp; English.

CC The invention relates to the complete nucleic acid sequence of a new
 CC human hepatitis B virus (HBV) genotype, provisionally named genotype
 CC G. This genotype was found with a high prevalence in patients
 CC chronically infected with HBV and residing in Europe and the USA. The
 CC invention relates to a fully defined sequence of 3248 nucleotides as
 CC given in specification, a sequence with 92% identity to the given
 CC sequence, or sequence that is degenerate to the mentioned sequences.
 CC These polynucleotides are useful for HBV genotyping. The proteins
 CC encoded by the polynucleotides are useful for detecting antibodies in
 CC a biological sample. Ligands that bind to the proteins and antibodies
 CC and for detecting HBcAg and HBeAg (precore precursor proteins). They
 CC are also useful for preparing a vaccine or medicament for treating
 CC HBV infections. The present sequence is provided in an amino acid
 CC sequence alignment of the preCore and Core regions of the different
 CC HBV genotypes.

SQ Sequence 212 AA;

Query Match 90.4%; Score 1037.5; DB 22; Length 212;
 Best Local Similarity 96.6%; Pred. No. 5e-98;
 Matches 196; Conservative 1; Mismatches 1; Indels 5; Gaps 1;

QY 13 PTVQASKLCLEGLMGNDIDPYKFGATVELLSFLSDPFPVSVDLDTASALYREALBSP 72
 DB 15 PTVQASKLCLEGLMGNDIDPYKFGATVELLSFLSDPFPVSVDLDTASALYREALBSP 74
 QY 73 EHCSPHHTALROAILCQGLMTLTA TWGVNLEDPFRGASDLVSVYNTNMGLKFRQL 132
 DB 75 EHCSPHHTALROAILCQGLMTLTA TWGVNLEDP-----ASRDLVSVYNTNMGLKFRQL 129
 QY 133 LMFHISCLTFGRRETVIEYLVSPGVWIRTPPAYRPNAPILSTLPETTVYRRGRSPRRRT 192
 DB 130 LMFHISCLTFGRRETVIEYLVSPGVWIRTPPAYRPNAPILSTLPETTVYRRGRSPRRRT 189
 QY 193 PSPRRRSQSPRRRSQSPRRSQ 215
 DB 190 PSPRRRSQSPRRRSQSPRRSQ 212

RESULT 8

AA080914
 ID AA080914 standard; Protein; 212 AA.

AC AA080914;

DT 09-APR-2002 (first entry)

DE Hepatitis B virus core antigen variant (HBcAg) #19.

KW Vaccine; molecular scaffold; pilus; pilin; HBcAg; antigen;
 KW hepatitis B virus capsid protein; JUN; FOS; HIV gp140;
 KW measles virus N protein; bee venom phospholipase; Th type 2 T-helper;
 KW T12; Sindis virus E2 protein; amyloid beta; influenza M2 antigen;
 KW human immunodeficiency virus infection; viral hepatitis; measles;
 KW chicken pox; pneumonia; tuberculosis; syphilis; malaria; allergy;
 KW cancer; chronic disease; arthritis; colitis; diabetes;
 KW multiple sclerosis.

OS Hepatitis B virus.

FN WO200185208-A2.

PD 15-NOV-2001.

PF 02-MAY-2001; 2001WO-IB00741.

PR 05-MAY-2000; 2000US-202341P.

PA (CYTO-) CYTOS BIOTECHNOLOGY AG.

PA (SEBB/) SEBBEL P.

PA (DUNA/) DUNANT N.

PA (BACH/) BACHMANN M.

PA (TIS/) TISSOT A.

PA (LECH/) LECHNER F.

PI Sebbel P, Dunant N, Bachmann M, Tissot A, Lechner F;

DR WPI; 2002-055561/07.

PT New composition, useful for vaccine production, comprises antigen or

PT antigenic determinant and non-natural molecular scaffold comprising

PT organizer and core particle such as bacterial pilus or pilin protein

PT -s

PS Disclosure; Page 222; 287pp; English.

CC The invention relates to a composition comprising: (a) a non-natural
 CC molecular scaffold (molecular scaffold) which comprises a core
 CC particle such as a bacterial pilus or pilin protein, a recombinant form
 CC of the protein, a virus-like particle or a hepatitis B virus capsid
 CC protein (HBcAg), and an organizer; and (b) an antigen or antigenic
 CC determinant, where the molecular scaffold and antigenic determinant
 CC interact to form an ordered and repetitive antigen array. Suitable
 CC antigenic determinants include UN, FOS, HIV gp140, measles virus N
 CC protein, bee venom phospholipase, Sindis virus E2 protein, amyloid beta

CC derived peptides and influenza M2 antigen. The composition (or vaccine)
 CC is useful for immunisation, by administration to a subject, where the
 CC administration produces an immune response, such as humoral, cellular or
 CC protective immune response, preferably a Th type 2 T-helper (Th2)
 CC response that is specific for the antigenic determinant. The
 CC administration induces antibodies specific for the antigenic determinant
 CC of a subtype corresponding to the Th2 subtype in the subject. The subject
 CC does not generate a Th2 subtype that is specific for plus or pili for the
 CC polypeptide or vaccine determinant. The composition is useful for the
 CC production of vaccines for prevention of infectious diseases such as
 CC human immunodeficiency virus, viral hepatitis, measles, chicken pox,
 CC pneumonia, tuberculosis, syphilis, colitis, diabetes and multiple
 CC cancer, and chronic diseases induced or accelerated by a Th1 type
 CC immune response, such as arthritis, colitis, diabetes and multiple
 CC sclerosis. The composition is useful to generate defined self-specific
 CC antibodies and specific immune responses of the Th2 type and allows the
 CC creation of highly efficient vaccines against infectious diseases, and
 CC for treating allergy, cancer, and chronic diseases induced or accelerated
 CC by a Th1 type immune response. The present sequence is a peptide or
 CC protein incorporated into the compositions of the invention.

XX Sequence 212 AA:

Query Match 90.4%; Score 1037.5; DB 23; Length 212;
 Best Local Similarity 96.6%; Pred. No. 5e-98; Indels 5; Gaps 1;
 Matches 196; Conservative 1; Mismatches 1;

QY 13 PTVOASKLCIGWLMGMDIDPYKFGATVLLSFLPSDFPPSYRDLDTASALYREALSPP 72
 DB 15 PTVOASKLCIGWLMGMDIDPYKFGATVLLSFLPSDFPPSYRDLDTASALYREALSPP 74
 QY 73 EHCSPHHTALROAILCWGLMTLATVGVNLDPERRGASRDLYVSYNTNMGLKFRQL 132
 DB 75 EHCSPHHTALROAILCWGLMTLATVGVNLDPERRGASRDLYVSYNTNMGLKFRQL 129
 QY 133 LWFHISCLTGFRETVLELVSGFWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRT 192
 DB 130 LWFHISCLTGFRETVLELVSGFWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRT 189
 QY 193 PSPRRRSQSPRRRSQSPRRRSQSPRRRSQSPRRRSQSPRRRSQSPRRRSQSPRRRSQ 215
 DB 190 PSPRRRSQSPRRRSQSPRRRSQSPRRRSQSPRRRSQSPRRRSQSPRRRSQSPRRRSQ 212

RESULT 9
 AAU80915
 ID AAU80915 standard; Protein; 212 AA.

AC AAU80915;
 DT 09-APR-2002 (first entry)
 XX
 DE Hepatitis B virus core antigen variant (HBCAg) #20.

XX
 KW Vaccine; molecular scaffold; pilus; pili; HBCAg; antigen;
 KW hepatitis B virus capsid protein; JUN; FOS; HIV gp140;
 KW measles virus N protein; bee venom phospholipase; Th type 2 T-helper;
 KW Th1; Simbia virus E2 protein; amyloid beta; influenza M2 antigen;
 KW human immunodeficiency virus infection; viral hepatitis; measles;
 KW chicken pox; pneumonia; tuberculosis; syphilis; malaria; allergy;
 KW cancer; chronic disease; arthritis; colitis; diabetes;
 KW multiple sclerosis.

KW Hepatitis B virus.

OS WO200185208-A2.

PD 15-NOV-2001.

PF 02-MAY-2001; 2001WO-1B00741.

PR 05-MAY-2000; 2000US-202341P.

XX

PA (CYTO-) CYTOS BIOTECHNOLOGY AG.
 PA (SEBEL) SEBEL P.
 PA (DUNA) DUNANT N.
 PA (BACH) BACHMANN M.
 PA (TISS) TISSOT A.
 PA (LECH) LECHNER F.
 PI Sebel P, Dunant N, Bachmann M, Tissot A, Lechner F;
 DR WPI: 2002-055561/07.
 XX
 PT New composition, useful for vaccine production, comprising antigen or
 PT antigenic determinant and non-natural molecular scaffold comprising
 PT organizer and core particle such as bacterial pilus or pili protein
 PT -8
 PS Disclosure; Page 223; 287pp; English.

XX The invention relates to a composition comprising: (a) a non-natural
 CC molecular scaffold (molecular scaffold) which comprises a core
 CC particle such as a bacterial pilus or pili protein, a recombinant form
 CC of the protein, a virus-like particle or a hepatitis B virus capsid
 CC protein (HBCAg), and an organizer; and (b) an antigen or antigenic
 CC determinant, where the molecular scaffold and antigenic determinant
 CC interact to form an ordered and repetitive antigen array. Suitable
 CC antigenic determinants include JUN, FOS, HIV gp140, measles virus N
 CC protein, bee venom phospholipase, Simbia virus E2 protein, amyloid beta
 CC derived peptides and influenza M2 antigen. The composition (or vaccine)
 CC is useful for immunisation, by administration to a subject, where the
 CC administration produces an immune response, such as humoral, cellular or
 CC protective immune response, preferably a Th type 2 T-helper (Th2)
 CC response that is specific for the antigenic determinant. The
 CC administration induces antibodies specific for the antigenic determinant
 CC of a subtype corresponding to the Th2 subtype in the subject. The subject
 CC does not generate a Th2 subtype that is specific for plus or pili for the
 CC polypeptide or vaccine determinant. The composition is useful for the
 CC production of vaccines for prevention of infectious diseases such as
 CC human immunodeficiency virus, viral hepatitis, measles, chicken pox,
 CC pneumonia, tuberculosis, syphilis, colitis, diabetes and multiple
 CC cancer, and chronic diseases induced or accelerated by a Th1 type
 CC immune response, such as arthritis, colitis, diabetes and multiple
 CC sclerosis. The composition is useful to generate defined self-specific
 CC antibodies and specific immune responses of the Th2 type and allows the
 CC creation of highly efficient vaccines against infectious diseases, and
 CC for treating allergy, cancer, and chronic diseases induced or accelerated
 CC by a Th1 type immune response. The present sequence is a peptide or
 CC protein incorporated into the compositions of the invention.

XX Sequence 212 AA:

Query Match 90.4%; Score 1037.5; DB 23; Length 212;
 Best Local Similarity 96.6%; Pred. No. 5e-98; Indels 5; Gaps 1;
 Matches 196; Conservative 1; Mismatches 1;

QY 13 PTVOASKLCIGWLMGMDIDPYKFGATVLLSFLPSDFPPSYRDLDTASALYREALSPP 72
 DB 15 PTVOASKLCIGWLMGMDIDPYKFGATVLLSFLPSDFPPSYRDLDTASALYREALSPP 74
 QY 73 EHCSPHHTALROAILCWGLMTLATVGVNLDPERRGASRDLYVSYNTNMGLKFRQL 132
 DB 75 EHCSPHHTALROAILCWGLMTLATVGVNLDPERRGASRDLYVSYNTNMGLKFRQL 129
 QY 133 LWFHISCLTGFRETVLELVSGFWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRT 192
 DB 130 LWFHISCLTGFRETVLELVSGFWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRT 189
 QY 193 PSPRRRSQSPRRRSQSPRRRSQSPRRRSQSPRRRSQSPRRRSQSPRRRSQSPRRRSQ 215
 DB 190 PSPRRRSQSPRRRSQSPRRRSQSPRRRSQSPRRRSQSPRRRSQSPRRRSQSPRRRSQ 212

RESULT 10
 AAU80916

ID AAU80916 standard; Protein: 212 AA.
 XX AAU80916;
 AC
 DT 09-APR-2002 (first entry)
 DE Hepatitis B virus core antigen variant (HBCAg) #21.
 XX
 KW Vaccine; molecular scaffold; pilus; pilin; HBCAg; antigen;
 KW hepatitis B virus capsid protein; JUN; FOS; HIV gp140;
 KW measles virus N protein; bee venom phospholipase; Th type 2 T-helper;
 KW Th2; Sindbis virus E2 protein; amyloid beta; influenza M2 antigen;
 KW human immunodeficiency virus infection; viral hepatitis; measles;
 KW chicken pox; pneumonia; tuberculosis; syphilis; malaria; allergy;
 KW cancer; chronic disease; arthritis; colitis; diabetes;
 KW multiple sclerosis.
 XX
 OS Hepatitis B virus.
 XX
 PN MO200185208-A2.
 XX
 PD 15-NOV-2001.
 XX
 PF 02-MAY-2001; 2001WO-1B00741.
 XX
 PR 05-MAY-2000; 2000US-202341P.
 XX
 PA (CYTO-) CYTOS BIOTECHNOLOGY AG.
 PA (SEBB/) SEBBEL P.
 PA (DUNA/) DUNANT N.
 PA (BACH/) BACHMANN M.
 PA (TISS/) TISSOT A.
 PA (LECH/) LECHNER F.
 XX
 PI Sebbel P, Dunant N, Bachmann M, Tissot A, Lechner F;
 DR WPI; 2002-055561/07.
 PT New composition, useful for vaccine production, comprises antigen or
 PT antigenic determinant and non-natural molecular scaffold comprising
 PT organizer and core particle such as bacterial pilus or pilin protein
 PT -8
 XX
 PS Disclosure; Page 223-224; 287pp; English.
 XX
 CC The invention relates to a composition comprising: (a) a non-natural
 CC molecular scaffold (molecular scaffold) which comprises a core
 CC particle such as a bacterial pilus or pilin protein, a recombinant form
 CC of the protein, a virus-like particle or a hepatitis B virus capsid
 CC protein (HBCAg), and an organizer; and (b) an antigen or antigenic
 CC determinant, where the molecular scaffold and antigenic determinant
 CC interact to form an ordered and repetitive antigen array. Suitable
 CC antigenic determinants include JUN, FOS, HIV gp140, measles virus N
 CC protein, bee venom phospholipase, Sindbis virus E2 protein, amyloid beta
 CC derived peptides and influenza M2 antigen. The composition (or vaccine)
 CC is useful for immunisation, by administration to a subject, where the
 CC administration produces an immune response, such as humoral, cellular or
 CC protective immune response, preferably a Th type 2 T-helper (Th2)
 CC response that is specific for the antigenic determinant. The
 CC administration induces antibodies specific for the antigenic determinant
 CC of a subtype corresponding to the Th2 subtype in the subject. The subject
 CC does not generate a Th2 subtype that is specific for pilus or pilin
 CC polypeptide or antigenic determinant. The composition is useful for the
 CC production of vaccines for prevention of infectious diseases such as
 CC human immunodeficiency virus, viral hepatitis, measles, chicken pox,
 CC pneumonia, tuberculosis, syphilis, malaria, and for treating allergy,
 CC cancer, and chronic diseases induced or accelerated by a Th1 type
 CC immune response, such as arthritis, colitis, diabetes and multiple
 CC sclerosis. The composition is useful to generate defined self-specific
 CC antibodies and specific immune responses of the Th2 type and allows the
 CC creation of highly efficient vaccines against infectious diseases, and
 CC for treating allergy, cancer, and chronic diseases induced or accelerated
 CC by a Th1 type immune response. The present sequence is a peptide or

CC protein incorporated into the compositions of the invention.
 XX
 XX Sequence 212 AA;
 SQ
 Query Match 90.4%; Score 1037.5; DB 23; Length 212;
 Best Local Similarity 96.6%; Pred. No. 5e-98;
 Matches 196; Conservative 1; Mismatches 1; Indels 5; Gaps 1;
 QY 13 PTVOASKCLGLMWMGDIDPYKBFATVETLSFLPSDFPSVVDLDTASALYREALSP 72
 DB 15 PTVOASKCLGLMWMGDIDPYKBFATVETLSFLPSDFPSVVDLDTASALYREALSP 74
 QY 73 BHCSPHHTLRRAIICWGHMTLATWGVNLDPDFRQDASHLVSVYNTNMGLKFRQL 132
 DB 75 BHCSPHHTLRRAIICWGHMTLATWGVNLDP-----ASRDLVSVYNTNMGLKFRQL 129
 QY 133 IMFHISCLTFGRVETVETLVSPGVWIRTPPAVRPNAPILSTLPTTVRRGRSPRRRT 192
 DB 130 IMFHISCLTFGRVETVETLVAFGVWIRTPPAVRPNAPILSTLPTTVRRGRSPRRRT 189
 QY 193 PSFRRRRSQSPRRRRSQSPREPQC 215
 DB 190 PSFRRRRSQSPRRRRSQSPREPQC 212

RESULT 11
 ID AAU80916 standard; Protein: 212 AA.
 XX AAU80916;
 AC
 DT 09-APR-2002 (first entry)
 DE Hepatitis B virus core antigen variant (HBCAg) #24.
 XX
 KW Vaccine; molecular scaffold; pilus; pilin; HBCAg; antigen;
 KW hepatitis B virus capsid protein; JUN; FOS; HIV gp140;
 KW measles virus N protein; bee venom phospholipase; Th type 2 T-helper;
 KW Th2; Sindbis virus E2 protein; amyloid beta; influenza M2 antigen;
 KW human immunodeficiency virus infection; viral hepatitis; measles;
 KW chicken pox; pneumonia; tuberculosis; syphilis; malaria; allergy;
 KW cancer; chronic disease; arthritis; colitis; diabetes;
 KW multiple sclerosis.
 XX
 OS Hepatitis B virus.
 XX
 PN MO200185208-A2.
 XX
 PD 15-NOV-2001.
 XX
 PF 02-MAY-2001; 2001WO-1B00741.
 XX
 PR 05-MAY-2000; 2000US-202341P.
 XX
 PA (CYTO-) CYTOS BIOTECHNOLOGY AG.
 PA (SEBB/) SEBBEL P.
 PA (DUNA/) DUNANT N.
 PA (BACH/) BACHMANN M.
 PA (TISS/) TISSOT A.
 PA (LECH/) LECHNER F.
 XX
 PI Sebbel P, Dunant N, Bachmann M, Tissot A, Lechner F;
 DR WPI; 2002-055561/07.
 PT New composition, useful for vaccine production, comprises antigen or
 PT antigenic determinant and non-natural molecular scaffold comprising
 PT organizer and core particle such as bacterial pilus or pilin protein
 PT -8
 XX
 PS Disclosure; Page 226; 287pp; English.
 XX
 CC The invention relates to a composition comprising: (a) a non-natural

QY 193 PPSRRRSQSPRRRSQSPRRQC 215
 DB 190 PPSRRRSQSPRRRSQSPRRQC 212

RESULT 13

AAU80922 standard; Protein; 212 AA.

AAU80922;

09-APR-2002 (first entry)

Hepatitis B virus core antigen variant (HBcAg) #27.

Vaccine; molecular scaffold; pilus; pili; HBcAg; antigen; hepatitis B virus capsid protein; JUN; FOS; HIV gp140; measles virus N protein; bee venom phospholipase; Th type 2 T-helper; Th2; Sindbis virus E2 protein; amyloid beta; influenza M2 antigen; human immunodeficiency virus infection; viral hepatitis; measles; chicken pox; pneumonia; tuberculosis; syphilis; malaria; allergy; cancer; chronic disease; arthritis; colitis; diabetes; multiple sclerosis.

Hepatitis B virus.

WO200185208-A2.

15-NOV-2001.

02-MAY-2001; 2001WO-1B00741.

05-MAY-2000; 2000US-202341P.

(CYTO-) CYTOS BIOTECHNOLOGY AG.

(SEBB/) SEBBEL P.

(DUNA/) DUNANT N.

(BACH/) BACHMANN M.

(TISS/) TISSOT A.

(LECH/) LECHNER F.

Sebbel P, Dunant N, Bachmann M, Tissot A, Lechner F,

WPI; 2002-055561/07.

New composition, useful for vaccine production, comprises antigen or antigenic determinant and non-natural molecular scaffold comprising organizer and core particle such as bacterial pilus or pilin protein

Disclosure; Page 228-229; 287pp; English.

The invention relates to a composition comprising: (a) a non-natural molecular scaffold (molecular scaffold) which comprises a core particle such as a bacterial pilus or pilin protein, a recombinant form of the protein, a virus-like particle or a hepatitis B virus capsid protein (HBcAg), and an organizer; and (b) an antigen or antigenic determinant, where the molecular scaffold and antigenic determinant interact to form an ordered and repetitive antigen array. Suitable antigenic determinants include JUN, FOS, HIV gp140, measles virus N protein, bee venom phospholipase, Sindbis virus E2 protein, amyloid beta derived peptides and influenza M2 antigen. The composition (or vaccine) is useful for immunisation, by administration to a subject, where the administration produces an immune response, such as humoral, cellular or response that is specific for the antigenic determinant. The administration induces antibodies specific for the antigenic determinant of a subtype corresponding to the Th2 subtype in the subject. The subject does not generate a Th2 subtype that is specific for pilus or pilin polypeptide or antigenic determinant. The composition is useful for the production of vaccines for prevention of infectious diseases such as human immunodeficiency virus, viral hepatitis, measles, chicken pox,

CC pneumonia, tuberculosis, syphilis, malaria, and for treating allergy, CC cancer, and chronic diseases induced or accelerated by a Th1 type CC immune response, such as arthritis, colitis, diabetes and multiple CC sclerosis. The composition is useful to generate defined self-specific CC antibodies and specific immune responses of the Th2 type and allows the CC creation of highly efficient vaccine against infectious diseases, and CC for treating allergy, cancer, and chronic diseases induced or accelerated by a Th1 type immune response. The present sequence is a peptide or protein incorporated into the compositions of the invention.

Sequence 212 AA;

Query Match 90.1%; Score 1034.5; DB 23; Length 212;

Best Local Similarity 96.6%; Pred. No. 1e-97;

Matches 196; Conservative 0; Mismatches 2; Indels 5; Gaps 1;

QY 13 PTVQASKLCIGLWGMDDIDPEYKFGATVYELSPSPDPSPVRLDIDPALSALYREALESP 72

DB 15 PTVQASKLCIGLWGMDDIDPEYKFGATVYELSPSPDPSPVRLDIDPALSALYREALESP 74

QY 73 HICSPHATLALQALICWGLMTLATVGVNLEDPEDFGDARDLVSVYVNTNMGILKFRQL 132

DB 75 HICSPHATLALQALICWGLMTLATVGVNLEDPEDFGDARDLVSVYVNTNMGILKFRQL 129

QY 133 LMFHISCLTPGRETIVYELVSGFWIRTPPAYRPPNAPILSTLETTYRRRGRSPRRRT 192

DB 130 LMFHISCLTPGRETIVYELVSGFWIRTPPAYRPPNAPILSTLETTYRRRGRSPRRRT 189

QY 193 PPSRRRSQSPRRRSQSPRRQC 215

DB 190 PPSRRRSQSPRRRSQSPRRQC 212

RESULT 14

AAU80923 standard; Protein; 212 AA.

AAU80923;

09-APR-2002 (first entry)

Hepatitis B virus core antigen variant (HBcAg) #28.

Vaccine; molecular scaffold; pilus; pili; HBcAg; antigen; hepatitis B virus capsid protein; JUN; FOS; HIV gp140; measles virus N protein; bee venom phospholipase; Th type 2 T-helper; Th2; Sindbis virus E2 protein; amyloid beta; influenza M2 antigen; human immunodeficiency virus infection; viral hepatitis; measles; chicken pox; pneumonia; tuberculosis; syphilis; malaria; allergy; cancer; chronic disease; arthritis; colitis; diabetes; multiple sclerosis.

Hepatitis B virus.

WO200185208-A2.

15-NOV-2001.

02-MAY-2001; 2001WO-1B00741.

05-MAY-2000; 2000US-202341P.

(CYTO-) CYTOS BIOTECHNOLOGY AG.

(SEBB/) SEBBEL P.

(DUNA/) DUNANT N.

(BACH/) BACHMANN M.

(TISS/) TISSOT A.

(LECH/) LECHNER F.

Sebbel P, Dunant N, Bachmann M, Tissot A, Lechner F,

WPI; 2002-055561/07.

PT New composition, useful for vaccine production, comprises antigen or
PT antigenic determinant and non-natural molecular scaffold comprising
PT organizer and core particle such as bacterial pilus or pilin protein
-6

PS Disclosure; Page 229-230; 287pp; English.

XX The invention relates to a composition comprising: (a) a non-natural
CC molecular scaffold (molecular scaffold) which comprises a core
CC particle such as a bacterial pilus or pilin protein, a recombinant form
CC of the protein, a virus-like particle or a hepatitis B virus capsid
CC protein (HBcAg), and an organizer; and (b) an antigen or antigenic
CC determinant, where the molecular scaffold and antigenic determinant
CC interact to form an ordered and repetitive antigen array. Suitable
CC antigenic determinants include JUN, FOS, HIV gp140, measles virus N
CC protein, bee venom phospholipase, Simble virus E2 protein, amyloid beta
CC derived peptides and influenza M2 antigen. The composition (or vaccine)
CC is useful for immunisation. By administration to a subject, where the
CC administration produces an immune response, such as humoral, cellular or
CC protective immune response, preferably a Th type 2 T-helper (Th2)
CC response that is specific for the antigenic determinant. The
CC administration induces antibodies specific for the antigenic determinant
CC of a subtype corresponding to the Th2 subtype in the subject. The subject
CC does not generate a Th2 subtype that is specific for pilus or pilin
CC polypeptide or antigenic determinant. The composition is useful for the
CC production of vaccines for prevention of infectious diseases such as
CC human immunodeficiency virus, viral hepatitis, measles, chicken pox,
CC pneumonia, tuberculosis, syphilis, malaria, and for treating allergy,
CC cancer, and chronic diseases induced or accelerated by a Th1 type
CC immune response, such as arthritis, colitis, diabetes and multiple
CC sclerosis. The composition is useful to generate defined self-specific
CC antibodies and specific immune responses of the Th2 type and allows the
CC creation of highly efficient vaccines against infectious diseases, and
CC for treating allergy, cancer, and chronic diseases induced or accelerated
CC by a Th1 type immune response. The present sequence is a peptide or
CC protein incorporated into the compositions of the invention.

XX Sequence 212 AA;

XX Query Match 90.1%; Score 1034.5; DB 23; Length 212;
XX Best Local Similarity 96.6%; Pred. No. 1e-97;
XX Matches 196; Conservative 0; Mismatches 2; Indels 5; Gaps 1;

QY 13 PTVOASKLCIGLMGMDIDPYKFGATVVLSPSDPSPVVDLDTASALYREALSP 72
DB 15 PTVOASKLCIGLMGMDIDPYKFGATVVLSPSDPSPVVDLDTASALYREALSP 74
QY 73 EHSPHHTALROAILCMGELMTLATWGVNLBDEPFRDASRDLYVSYVTNMGKFRQL 132
DB 75 EHSPHHTALROAILCMGELMTLATWGVNLBDEPFRDASRDLYVSYVTNMGKFRQL 129
QY 133 LWFHISCLTGFRETIVLYVSGVWIRTPPAYRPNAPILSTLETTVVRGRSPRRRT 192
DB 130 LWFHISCLTGFRETIVLYVSGVWIRTPPAYRPNAPILSTLETTVVRGRSPRRRT 189
QY 193 PSPRRRSQSPRRRSQSPRRPQC 215
DB 190 PSPRRRSQSPRRRSQSPRRPQC 212

RESULT 15
AAU80924 standard; Protein; 212 AA.
ID AAU80924 standard; Protein; 212 AA.
XX
XX AAU80924;
AC
XX
XX 09-APR-2002 (first entry)
DT
XX
XX Hepatitis B virus core antigen variant (HBcAg) #29.
DE
XX
XX Vaccine; molecular scaffold; pilus; pilin; HBcAg; antigen;
KM hepatitis B virus capsid protein; JUN; FOS; HIV gp140;
KM measles virus N protein; bee venom phospholipase; Th type 2 T-helper;

KM Th2; Simble virus E2 protein; amyloid beta; influenza M2 antigen;
KM human immunodeficiency virus infection; viral hepatitis; measles;
KM chicken pox; pneumonia; tuberculosis; syphilis; malaria; allergy;
KM cancer; chronic disease; arthritis; colitis; diabetes;
KM multiple sclerosis.

OS Hepatitis B virus.

XX WO200185208-A2.

XX 15-NOV-2001.

XX 02-MAY-2001; 2001WO-1B00741.

XX 05-MAY-2000; 2000US-202341P.

XX (CYTO-) CYTOS BIOTECHNOLOGY AG.

XX (SEBB/) SEBBEL P.

XX (DUNA/) DUNANT N.

XX (BACH/) BACHMANN M.

XX (TISS/) TISSOT A.

XX (LECH/) LECHNER F.

XX Sebbel P, Dunant N, Bachmann M, Tissot A, Lechner F;

XX WPI; 2002-055561/07.

XX New composition, useful for vaccine production, comprises antigen or

XX antigenic determinant and non-natural molecular scaffold comprising

XX organizer and core particle such as bacterial pilus or pilin protein

XX -6

XX Disclosure; Page 230; 287pp; English.

XX The invention relates to a composition comprising: (a) a non-natural
CC molecular scaffold (molecular scaffold) which comprises a core
CC particle such as a bacterial pilus or pilin protein, a recombinant form
CC of the protein, a virus-like particle or a hepatitis B virus capsid
CC protein (HBcAg), and an organizer; and (b) an antigen or antigenic
CC determinant, where the molecular scaffold and antigenic determinant
CC interact to form an ordered and repetitive antigen array. Suitable
CC antigenic determinants include JUN, FOS, HIV gp140, measles virus N
CC protein, bee venom phospholipase, Simble virus E2 protein, amyloid beta
CC derived peptides and influenza M2 antigen. The composition (or vaccine)
CC is useful for immunisation. By administration to a subject, where the
CC administration produces an immune response, such as humoral, cellular or
CC protective immune response, preferably a Th type 2 T-helper (Th2)
CC response that is specific for the antigenic determinant. The
CC administration induces antibodies specific for the antigenic determinant
CC of a subtype corresponding to the Th2 subtype in the subject. The subject
CC does not generate a Th2 subtype that is specific for pilus or pilin
CC polypeptide or antigenic determinant. The composition is useful for the
CC production of vaccines for prevention of infectious diseases such as
CC human immunodeficiency virus, viral hepatitis, measles, chicken pox,
CC pneumonia, tuberculosis, syphilis, malaria, and for treating allergy,
CC cancer, and chronic diseases induced or accelerated by a Th1 type
CC immune response, such as arthritis, colitis, diabetes and multiple
CC sclerosis. The composition is useful to generate defined self-specific
CC antibodies and specific immune responses of the Th2 type and allows the
CC creation of highly efficient vaccines against infectious diseases, and
CC for treating allergy, cancer, and chronic diseases induced or accelerated
CC by a Th1 type immune response. The present sequence is a peptide or
CC protein incorporated into the compositions of the invention.

XX Sequence 212 AA;

XX Query Match 90.1%; Score 1034.5; DB 23; Length 212;
XX Best Local Similarity 96.1%; Pred. No. 1e-97; 1; Indels 5; Gaps 1;
XX Matches 195; Conservative 2; Mismatches 1;

QY 13 PTVOASKLCIGLMGMDIDPYKFGATVVLSPSDPSPVVDLDTASALYREALSP 72
DB 15 PTVOASKLCIGLMGMDIDPYKFGATVVLSPSDPSPVVDLDTASALYREALSP 74

Qy 73 EHCSPHHTALRQAILCWEGLMTLATWVGVLBDPEFRGDASRDVVSYYNTNMGKPROL 132
Db 75 EHCSPHHTALRQAILCWEGLMTLATWVGVLBDPEFRGDASRDVVSYYNTNMGKPROL 129
Qy 133 LMFHISCLTREGRETVLELVSGVWIRTPPAYRPNAPILSTLDETTVVRGRSPRRRT 192
Db 130 LMFHISCLTREGRETVLELVSGVWIRTPPAYRPNAPILSTLDETTVVRGRSPRRRT 189
Qy 193 PSPRRRRSOSPRRRRSOSPQ 215
Db 190 PSPRRRRSOSPRRRRSOSPQ 212

Search completed: February 3, 2003, 09:23:51
Job time : 35.459 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 09:16:26 ; Search time 0.439024 Seconds

(without alignments)
1133.686 Million cell updates/sec

Title: US-09-890-752a-20
Perfect score: 61
Sequence: 1 PLSSTFSRIGDP 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61	100.0	389	1 VMSA_HPBVY	P03138 hepatitis b
2	59	96.7	389	1 VMSA_HPBVA	P24025 hepatitis b
3	59	96.7	389	1 VMSA_HPBVA	P03139 hepatitis b
4	54	88.5	400	1 VMSA_HPBVA	P03140 hepatitis b
5	46	75.4	400	1 VMSA_HPBVA	P12934 hepatitis b
6	44	72.1	400	1 VMSA_HPBVA	P02317 hepatitis b
7	40	65.6	64	1 ED02_HUMAN	O15263 homo sapien
8	40	65.6	389	1 VMSA_HPBVA	P12911 hepatitis b
9	38	62.3	389	1 VMSA_HPBVA	P17398 hepatitis b
10	38	62.3	389	1 VMSA_HPBVA	P17399 hepatitis b
11	38	62.3	537	1 AREH_SCHRO	Q10269 schistosom
12	38	62.3	1786	1 VIRA_CHLTR	O84337 chlamydia t
13	37	60.7	495	1 YOR2_CABEL	O09307 caenorhabd
14	37	60.7	533	1 VWIN_SYNZ3	O55179 chlamydia m
15	36	59.0	334	1 DCUP_CHLMU	O90147 chlamydia m
16	36	59.0	389	1 VMSA_HPBVA	P03142 hepatitis b
17	36	59.0	400	1 VMSA_HPBVA	P03141 hepatitis b
18	36	59.0	400	1 VMSA_HPBVA	P17101 hepatitis b
19	36	59.0	593	1 ME44_MCV1	Q88224 molluscum c
20	36	59.0	695	1 EFGI_SYNZ3	P28371 mycobacteri
21	36	59.0	701	1 EFGI_MICU1	P09952 micrococcu
22	36	59.0	1081	1 PDR6_YEAST	P32767 saccharomyc
23	35	57.4	341	1 EFG_STRPA	P29541 streptomyce
24	35	57.4	389	1 VMSA_HPBVA	P17397 hepatitis b
25	35	57.4	400	1 VMSA_HPBVA	Q05496 hepatitis b
26	35	57.4	538	1 IFGD_SHIFL	Q07566 shigella fl
27	35	57.4	701	1 EFG_MYCLE	P30767 mycobacteri
28	35	57.4	701	1 EFG_MYCTU	O53790 mycobacteri
29	35	57.4	708	1 EFGI_STRCO	P40173 streptomyce
30	35	57.4	1302	1 RRPL_BT110	P13840 bluecorgue
31	34	55.7	70	1 NTSCG_TIBAS	P36262 lybacteri
32	34	55.7	225	1 NTCA_SYNZ3	P33779 synchocyst
33	34	55.7	247	1 YGPF_ECOLI	P52037 escherichia

34	34	55.7	250	1 VGLL_HSV62	P52526 human herpe
35	34	55.7	260	1 NMA_HUMAN	Q13145 homo sapien
36	34	55.7	265	1 CB23_TOBAC	P27494 nicotiana t
37	34	55.7	268	1 CB23_HORVU	P27523 hordeum vul
38	34	55.7	270	1 CB12_LYCER	P10708 lycopersico
39	34	55.7	270	1 CB12_PERTY	P13869 petunia hyb
40	34	55.7	272	1 MURI_MYCLE	P18605 mycobacteri
41	34	55.7	283	1 LGT_HELPV	O92kp6 hellicobacte
42	34	55.7	284	1 LGT_HELPV	O25609 hellicobacte
43	34	55.7	312	1 PHO4_YEAST	P07270 saccharomyc
44	34	55.7	344	1 CDK7_CARAU	P51953 carassius a
45	34	55.7	421	1 PHFL_DESVH	P07598 deeslovibir

ALIGNMENTS

RESULT 1	VMSA_HPBVY	STANDARD:	PRT:	389 AA.
ID	P03138:			
AC	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	01-NOV-1997 (Rel. 35, Last annotation update)			
DE	Major surface antigen precursor.			
OS	Hepatitis B virus (subtype ayw).			
OC	Vitruvius, Retrovird viruses; Hepadnaviridae; Orthohepadnavirus.			
OK	NCBI_TaxID=10418;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=81012091; PubMed=399327;			
RA	Gallibert F., Mandart E., Pitoussi F., Tiollais P., Charnay P.;			
RT	"Nucleotide sequence of the hepatitis B virus genome (subtype ayw)			
RT	cloned in E. coli."			
RL	Nature 281:646-650(1979).			
RN	[2]			
RP	SEQUENCE FROM N.A. (CLONE PHB320).			
RX	MEDLINE=85204397; PubMed=396597;			
RA	Bischo V., Pushko P., Dreiling D., Pumpen P., Gren E.Y.;			
RT	"Subtype ayw variant of hepatitis B virus. DNA primary structure			
RT	analysis."			
RN	FEBS Lett. 185:208-212(1985).			
RP	[3]			
RA	SEQUENCE FROM N.A.			
RA	Pumpen P., Kozlovskaya T.M., Borisova G.L., Byuchko V.V.;			
RA	Dishier A.V., Kallio Y.V., Pudova N.V., Gren E.Y., Krievlana V.Y.;			
RT	Kutakin R.A.;			
RT	"Synthesis of the surface antigen of hepatitis B virus in Escherichia			
RL	coli."			
RL	Dokl. Biochem. 271:246-249(1984).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sdb.ch/announce/			
CC	or send an email to license@isb-sdb.ch).			
CC	-----			
DR	EMBL; V01460; -; NOT ANNOTATED CDS.			
DR	EMBL; X02496; CAA26324.1; ALT INIT.			
DR	EMBL; M12393; AAA45496.1; ALT INIT.			
DR	PIR; A03703; SAVLH.			
DR	PIR; A03704; SAVLH.			
DR	InterPro: IPR000349; Hepvir_surf.			
DR	Pfam: PF00695; VMSA; 1.			
KW	Antigen.			
FT	PROPEP	1	163	
FT	CHAIN	164	389	
FT	CARBOHYD	4	4	
FT	CARBOHYD	112	112	
FT	CARBOHYD	166	166	
MAJOR SURFACE ANTIGEN.				
N-LINKED (GLCNAC. . .)				(POTENTIAL).
N-LINKED (GLCNAC. . .)				(POTENTIAL).
N-LINKED (GLCNAC. . .)				(POTENTIAL).

```

FT CARBOHYD 309 309 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 75 75 Q -> E (IN REF. 2).
FT CONFLICT 147 147 A -> S (IN REF. 2).
FT CONFLICT 150 150 L -> I (IN REF. 2).
FT CONFLICT 288 290 MTT -> TTP (IN REF. 2 AND 3).
SQ SEQUENCE 389 AA; 42766 MM; 6DC9E682DA694F63 CRC64;

Query Match 100.0%; Score 61; DB 1; Length 389;
Best Local Similarity 100.0%; Pred. No. 0.00052;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSISFSTRIGDP 12
Db 149 PLSISFSTRIGDP 160

RESULT 2
VMSA_HPBVA STANDARD; PRT; 389 AA.
ID VMSA_HPBVA STANDARD; PRT; 389 AA.
AC P24025;
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-MAR-1992 (Rel. 21, Last annotation update)
DE Major surface antigen precursor.
OS Hepatitis B virus (strain alpha1).
OC Virusess; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
NC NCB1_TaxID=10411;
RN NCB1_TaxID=10411;
RX SEQUENCE FROM N.A.
RA MEDLINE=90266476; PubMed=2345966;
RA Tong S., Li J., Valtitsaki L., Tsepo C.;
RA "Active hepatitis B virus variants containing an inactive pre-C region.";
RT Virology 176:596-603(1990).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M3138; AAA45502.1; .
CC DR PIR; B34773; SAVLAI.
CC DR InterPro; IPR000349; Hepv1r_surfag.
CC DR Pfam; PF00695; VMSA; 1.
CC KW Antigen.
FT PROPEP 1 163 MAJOR SURFACE ANTIGEN.
FT CHAIN 164 389 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4 4 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 309 309 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 389 AA; 42733 MM; B806D0D42B23BF5 CRC64;

Query Match 96.7%; Score 59; DB 1; Length 389;
Best Local Similarity 91.7%; Pred. No. 0.0012;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSISFSTRIGDP 12
Db 149 PLSISFSTRIGDP 160

RESULT 3
VMSA_HPBVA STANDARD; PRT; 389 AA.
ID VMSA_HPBVA STANDARD; PRT; 389 AA.
AC P03139;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-OCT-1989 (Rel. 12, Last annotation update)

```

```

DE Major surface antigen precursor.
OS Hepatitis B virus (subtype adw).
OC Virusess; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
NC NCB1_TaxID=10419;
RN NCB1_TaxID=10419;
RX SEQUENCE FROM N.A.
RA MEDLINE=8102115; PubMed=399329;
RA Pasek W., Goto T., Gilbert W., Zink B., Scheller H., Mackay P.,
RA Leadbetter G., Murray K.;
RA "Hepatitis B virus genes and their expression in E. coli.";
RT Nature 282:575-579(1979).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J02202; AAA45487.1; ALT_INIT.
CC DR PIR; A93217; SAVLAI.
CC DR InterPro; IPR000349; Hepv1r_surfag.
CC DR Pfam; PF00695; VMSA; 1.
CC KW Antigen.
FT PROPEP 1 163 MAJOR SURFACE ANTIGEN.
FT CHAIN 164 389 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4 4 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 309 309 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 389 AA; 42801 MM; 1069EF6B4679A669 CRC64;

Query Match 96.7%; Score 59; DB 1; Length 389;
Best Local Similarity 91.7%; Pred. No. 0.0012;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSISFSTRIGDP 12
Db 149 PLSISFSTRIGDP 160

RESULT 4
VMSA_HPBVA STANDARD; PRT; 400 AA.
ID VMSA_HPBVA STANDARD; PRT; 400 AA.
AC P03140;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Major surface antigen precursor.
OS Hepatitis B virus (subtype adr).
OC Virusess; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
NC NCB1_TaxID=106820;
RN NCB1_TaxID=106820;
RX SEQUENCE FROM N.A.
RA MEDLINE=83168919; PubMed=6300776;
RA Ono Y., Onda H., Sasada R., Igarashi K., Sugino Y., Nishioke K.;
RA "The complete nucleotide sequences of the cloned hepatitis B virus
RA DNA, subtype adr and adr.";
RT Nucleic Acids Res. 11:1747-1757(1983).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
CC or send an email to license@1sb-sib.ch).

DR EMBL; V00867; CA24234.1; ALT_INIT.
DR PIR; A03705; SAVLA.

DR InterPro; IPR000349; Hepvir_surfAg.
DR Pfam; PF00695; VMSA; 1.
KW Antigen.

FT PROPEP 1 174
FT CHAIN 175 400 MAJOR SURFACE ANTIGEN.
FT CARBOHYD 15 15 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 123 123 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 320 320 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 400 AA; 43685 MW; 16DB0B15C6CB697A CRC64;

Query Match
Best Local Similarity 88.5%; Score 54; DB 1; Length 400;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSSISFRIGDP 12
Db 160 PSSISFRIGDP 171

RESULT 5

VMSA_HPBV4
ID VMSA_HPBV4 STANDARD; PRT; 400 AA.

AC P12934; Q67860;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Major surface antigen precursor.

GN S.
OS Hepatitis B virus (subtype adr).
OC Viruses; Retrovird virusae; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10409;
RN [1]

RP MEDLINE=83246570; PubMed=6306594;
RA Fujiyama A., Miyenohara A., Nozaki C., Yoneyama T., Ohtomo N.,
RA Matsubara K.;
RT "Cloning and structural analyses of hepatitis B virus DNA, subtype
adr."
RL Nucleic Acids Res. 11:4601-4610(1983).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
CC or send an email to license@1sb-sib.ch).

DR EMBL; X01587; CA25747.1; ALT_INIT.
DR EMBL; X01587; CA25743.1; ALT_INIT.
DR InterPro; IPR000349; Hepvir_surfAg.
DR Pfam; PF00695; VMSA; 1.
KW Antigen.

FT PROPEP 1 174
FT CHAIN 175 400 MAJOR SURFACE ANTIGEN.
FT CARBOHYD 15 15 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 123 123 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 320 320 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 400 AA; 43542 MW; 01053BBD24053F9 CRC64;

Query Match
Best Local Similarity 75.4%; Score 46; DB 1; Length 400;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSSISFRIGDP 12
Db 160 PSSISFRIGDP 171

RESULT 6

VMSA_HPBVP
ID VMSA_HPBVP STANDARD; PRT; 400 AA.

AC 002317;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE Major surface antigen precursor.

GN S.
OS Hepatitis B virus (subtype adr / strain Philippine/PFDM294).
OC Viruses; Retrovird virusae; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=31514;
RN [1]

RP SEQUENCE FROM N.A.
RA Estacio R.C., Chavez C.C., Okamoto H., Lingao A.L., Reyes M.T.,
RA Domingo E., Mayumi M.,
RT "Nucleotide sequence of a hepatitis B virus genome of subtype adr
isolated from a Philippine: comparison with the reported three genomes
of the same subtype."
RL J. Gastroenterol. Hepatol. 3:215-222(1988).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
CC or send an email to license@1sb-sib.ch).

DR EMBL; M5763; AAA6860.1;
DR InterPro; IPR000349; Hepvir_surfAg.
DR Pfam; PF00695; VMSA; 1.
KW Antigen.

FT PROPEP 1 174
FT CHAIN 175 400 MAJOR SURFACE ANTIGEN.
FT CARBOHYD 15 15 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 123 123 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 320 320 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 400 AA; 43740 MW; 7EBD53A004136BD2 CRC64;

Query Match
Best Local Similarity 72.1%; Score 44; DB 1; Length 400;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PSSISFRIGDP 12
Db 161 PSSISFRIGDP 171

RESULT 7

ID BD02_HUMAN
ID BD02_HUMAN STANDARD; PRT; 64 AA.

AC O15253;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Beta-defensin 2 precursor (hBD-2) (Skin-antimicrobial peptide 1)
DE (SAP).
GN DEF82.

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutera; Primates; Carnivora; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=97345625; PubMed=9202117;

RA Harder J., Bartels J.H., Christophers E., Schroeder J.-M.;
 RT "A peptide antiproliferic from human skin.";
 RL Nature 387:861-861(1997).
 RP [2]
 RN SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=99051334; PubMed=9831658;
 RA Liu L., Wang L., Jia H.P., Zhao C., Heng H.H.Q., Schutte B.C.,
 RT McCray P.B. Jr., Ganz T.;
 RL "Structure and mapping of the human beta-defensin HBD-2 gene and its
 RT expression at sites of inflammation."; Gene 222:237-244(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20072673; PubMed=10603376;
 RA Diamond G., Kaiser V., Rhodes J., Russell J.P., Bevins C.L.;
 RT "Transcriptional regulation of beta-defensin gene expression in
 RL tracheal epithelial cells."; Infect. Immun. 68:113-119(2000).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Harder J., Schroeder J.-M.;
 RT "Transcriptional regulation of the human beta-defensin-2 (HBD-2).";
 RL Submitted (Aug-1997) to the EMBL/Genbank/DBJ databases.
 RN [5]
 RP SYNTHESIS OF 24-64.
 RX PubMed=12010514;
 RA Kluever E., Schulz A., Forssmann W.-G., Ademann K.;
 RT "Chemical synthesis of beta-defensins and LBP-1/hepcidin."; J.
 RL Pept. Res. 59:241-248(2002).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (1.35 ANGSTROMS).
 RX MEDLINE=20490730; PubMed=10906336;
 RA Hoover D.M., Rajshankar K.R., Blumenthal R., Puri A., Oppenheim J.J.,
 RT Chertov O., Lubkowsky J.;
 RL "The structure of human beta-defensin-2 shows evidence of higher
 RT order oligomerization."; J. Biol. Chem. 275:32911-32918(2000).
 RN [7]
 RP STRUCTURE BY NMR OF 28-64.
 RX MEDLINE=21571964; PubMed=11714914;
 RA Bauer F., Schweimer K., Kluever E., Conejo-Garcia J.R., Forssmann W.G.,
 RT Roach P., Ademann K., Slicht H.;
 RL "Structure determination of human and murine beta-defensins reveals
 RT structural conservation in the absence of significant sequence
 RT similarity"; Protein Sci. 10:2470-2479(2001).
 CC -1- FUNCTION: HAS ANTIBACTERIAL ACTIVITY (POTENTIAL).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE SKIN AND RESPIRATORY TRACT.
 CC -1- INDUCTION: BY INFLAMMATION.
 CC -1- SIMILARITY: BELONGS TO THE BETA-DEFENSIN FAMILY. LAP/TAP
 CC SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 CC EMBL; Z71389; CAA95992.1; -;
 DR EMBL; AF040153; AAC33549.1; -;
 DR EMBL; AF071216; AAC69554.1; -;
 DR EMBL; AJ000152; CAB65126.1; -;
 DR PDB; 1FD3; 01-NOV-00.
 DR PDB; 1FD4; 01-NOV-00.
 DR PDB; 1E4Q; 26-NOV-01.
 DR Genew; HGNC:2767; DEFB4.
 DR MIM; 602215; -;
 DR InterPro; IPR001855; Defensin_beta.
 DR InterPro; IPR001271; Defensin_mammal.

DR Pfam; PF00711; Defensin_beta; 1.
 DR SMART; SM00048; DEFSN; 1; structure.
 KW Antibiotic; Signal; 3D-structure.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 64 BETA-DEFENSIN 2.
 FT DISULFID 31 60
 FT DISULFID 38 53
 FT DISULFID 43 61
 SQ SEQUENCE 64 AA; 7038 MW; 05D6454CE7ACD10B CRC64;
 Query Match Score 40; DB 1; Length 64;
 Best Local Similarity 58.3%; Pred. No. 0.75;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1 PLPSIFSRIGDP 12
 Db 17 PLPSVFGIGDP 28
 RESULT 8
 ID VMSA_HPBVL STANDARD; PRT; 389 AA.
 AC P12911;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 01-APR-1990 (Rel. 14, Last annotation update)
 DE Major surface antigen precursor.
 GN S.
 OS Hepatitis B virus (strain 1bh / chimpanzee isolate).
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
 OX NCBI_Taxid=10414;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88258473; PubMed=2838576;
 RA Vaidin M., Wolstenholme A.J., Teiquaye K.N., Zuckerman A.J.,
 RA Harrison T.J.;
 RT "The complete nucleotide sequence of the genome of a hepatitis B
 RT virus isolated from a naturally infected chimpanzee."; J. Gen. Virol. 69:1383-1389(1988).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 CC EMBL; D00220; BAA00159.1; -;
 DR PIR; C28885; SAVLCP.
 DR InterPro; IPR000349; Hepvir_burFag.
 DR Pfam; PF00695; VMSA; 1.
 KW Antigen.
 FT PROPEP 1 163
 FT CHAIN 164 389
 FT CARBOHYD 4 4
 FT CARBOHYD 112 112
 FT CARBOHYD 166 166
 FT CARBOHYD 309 309
 SQ SEQUENCE 389 AA; 42539 MW; 87D1F3382A10DBD CRC64;
 Query Match Score 40; DB 1; Length 389;
 Best Local Similarity 63.6%; Pred. No. 4.9;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 2 LSGISFRIGDP 12
 Db 150 ISSVFSFTGDP 160
 RESULT 9
 ID VMSA_HPBVL STANDARD; PRT; 389 AA.
 VMSA_HPBVL


```

AC P17398;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DR Major surface antigen precursor.
CC S.
OS Hepatitis B virus (subtype adw / strain Japan/PTD233).
OC Viruses; Retroviral viruses; Hepadnaviridae; Orthohepadnavirus.
OK NCBI_TaxID=10413;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=89010694; PubMed=3171552;
RA Okamoto H., Tsuda F., Sakugawa H., Sastrosewignio R.I., Imai M.,
RA Miyakawa Y., Mayumi M.,
RT "Typing hepatitis B virus by homology in nucleotide sequence:
RT comparison of surface antigen subtypes."
RL J. Gen. Virol. 69:2575-2583 (1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib.ch).
CC -----
DR EMBL, D00329; NOT_ANNOTATED_CDS.
DR PIR, G28925; SAVLJ2.
DR InterPro, IPR000349; Hepvir_surfng.
DR Pfam, PF00695; VMSA; 1.
KW Antigen.
FT CHAIN 1 163 MAJOR SURFACE ANTIGEN.
FT PROPEP 1 389
FT CARBOHYD 164 4 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 309 309 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 389 AA; 42603 MW; 72F8D481924DC104 CRC64;

Query Match 62.3%; Score 38; DB 1; Length 389;
Best Local Similarity 63.6%; Pred. No. 12;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 LSISFRRIGDP 12
Db 150 ISSLSKRGDP 160

RESULT 10
VMSA_HPBVO STANDARD; PRT; 389 AA.
ID VMSA_HPBVO
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE Major surface antigen precursor.
GN S.
OS Hepatitis B virus (subtype adw / strain Okinawa/POD282).
OC Viruses; Retroviral viruses; Hepadnaviridae; Orthohepadnavirus.
OK NCBI_TaxID=10413;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=89010694; PubMed=3171552;
RA Okamoto H., Tsuda F., Sakugawa H., Sastrosewignio R.I., Imai M.,
RA Miyakawa Y., Mayumi M.,
RT "Typing hepatitis B virus by homology in nucleotide sequence:
RT comparison of surface antigen subtypes."
RL J. Gen. Virol. 69:2575-2583 (1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL, D00330; NOT_ANNOTATED_CDS.
DR PIR, H28925; SAVLJ2.
DR InterPro, IPR000349; Hepvir_surfng.
DR Pfam, PF00695; VMSA; 1.
KW Antigen.
FT CHAIN 1 163 MAJOR SURFACE ANTIGEN.
FT PROPEP 1 389
FT CARBOHYD 164 4 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 309 309 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 389 AA; 42408 MW; CECAC0DDA3B84A10 CRC64;

Query Match 62.3%; Score 38; DB 1; Length 389;
Best Local Similarity 63.6%; Pred. No. 12;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 LSISFRRIGDP 12
Db 150 ISSLSKRGDP 160

RESULT 11
AREH_SCHPO STANDARD; PRT; 537 AA.
ID AREH_SCHPO
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable sterol O-acetyltransferase (EC 2.3.1.26) (Sterol-ester
DE synthase).
GN SPAC1367.05.
OS Schizosaccharomyces pombe (Fission Yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OK NCBI_TaxID=4896;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Baeham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
RA Ruttenford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grymponprez B.,
RA Weltejn I., Vanstreels E., Rieger M., Schaefer M., Mueller-Nuer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Meisel D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leinach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Medler S., Gloux R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gauthier R., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garcon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cernutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Uesery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880 (2002).
CC -----
CC -1- CATALYTIC ACTIVITY: Acyl-CoA + cholesterol = CoA + cholesterol
CC ester.

```

```

CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC reticulum (by similarity).
CC -1- SIMILARITY: BELONGS TO THE STEROL O-ACYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.trib-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: Z69729; CA93593.1; -.
CC InterPro: IPR002688; ACAT.
CC Pfam: PF01800; ACAT. 1.
CC Hypothetical protein; Transferase; Transmembrane;
CC Endoplasmic reticulum; Acyltransferase.
CC TRANSMEM 98 118 POTENTIAL.
CC TRANSMEM 140 160 POTENTIAL.
CC TRANSMEM 174 194 POTENTIAL.
CC TRANSMEM 199 219 POTENTIAL.
CC TRANSMEM 305 325 POTENTIAL.
CC TRANSMEM 344 364 POTENTIAL.
CC TRANSMEM 384 404 POTENTIAL.
CC TRANSMEM 462 482 POTENTIAL.
CC TRANSMEM 488 508 POTENTIAL.
CC TRANSMEM 517 537 POTENTIAL.
CC CARBOHYD 250 250 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SEQUENCE 537 AA; 63098 MW; 41DBA79402B67A28 CRC64;
SQ
Query Match 62.3%; Score 38; DB 1; Length 537;
Best Local Similarity 58.3%; Pred. No. 16;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 PLSIFSRIGDP 12
DB 79 PKSLFDRVTD 90
RESULT 12
UVR_A CHLTR STANDARD; PRT; 1786 AA.
ID UVR_A CHLTR STANDARD; PRT; 1786 AA.
AC 084337;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Exonuclease ABC subunit A.
GN UVR_A OR CT333.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
(1)
RP SEQUENCE FROM N.A.
RC STRAIN=D/UM-3/Cx;
RX MEDLINE=99000809; PubMed=9784136;
RA Stephens R.S., Kaiman S., Lammell C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.;
RT "Genome sequence of an obligate intracellular pathogen of humans:
RT Chlamydia trachomatis."
RL Science 282:754-759(1998).
CC -1- FUNCTION: THE ABC EXCISION NUCLEASE IS A DNA REPAIR ENZYME THAT
CC CATALYZES THE EXCISION REACTION OF UV-DAMAGED NUCLEOTIDE SEGMENTS
CC PRODUCING OLIGOMERS HAVING THE MODIFIED BASE(S). UVR_A IS AN ATPASE
CC AND A DNA-BINDING PROTEIN THAT PREFERENTIALLY BINDS SINGLE-
CC STRANDED OR UV-IRRADIATED DOUBLE-STRANDED DNA (BY SIMILARITY).
CC -1- SUBUNIT: CONSISTS OF THREE SUBUNITS: UVR_A, UVR_B AND UVR_C.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity). UVR_A SUBFAMILY.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. UVR_A SUBFAMILY.
CC CONTAINS FOUR ABC DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.trib-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AE001306; AAC67928.1; -.
CC InterPro: IPR003439; ABC transporter.
CC InterPro: IPR004602; UVR_A.
CC Pfam: PD000005; ABC_tran; 2.
CC ProDom: PD000006; ABC_transport; 1.
CC TRIGRAMS: TIGR00630; uvrA. 1.
CC PROSITE: PS00211; ABC_TRANSPORTER; FALSE NEG.
CC SOS response; Excision nuclease; DNA repair; ATP-binding; Repeat;
CC DNA-binding; Zinc-finger; Complete proteome.
CC NP_BIND 32 39 ATP (POTENTIAL).
CC NP_BIND 625 632 ATP (POTENTIAL).
CC NP_BIND 964 971 ATP (POTENTIAL).
CC NP_BIND 1516 1523 ATP (POTENTIAL).
CC ZN_FING 719 742 C4-TYPE.
CC ZN_FING 1602 1628 C4-TYPE.
CC SEQUENCE 1786 AA; 196948 MW; 02D6862BE15DE070 CRC64;
SQ
Query Match 62.3%; Score 38; DB 1; Length 1786;
Best Local Similarity 63.6%; Pred. No. 57;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 2 LSLIFSRIGDP 12
DB 1042 LESLIFRIGHP 1052

```

```

RESULT 13
YOR2 CAEEL STANDARD; PRT; 495 AA.
ID YOR2 CAEEL STANDARD; PRT; 495 AA.
AC 009307;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 45.1 kDa protein F19C6.2 in chromosome X.
GN F19C6.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodermidae; Caenorhabditis.
OX NCBI_TaxID=6239;
(1)
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Harris B.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.trib-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: Z48006; CA88050.1; -.
CC WormPep: F19C6.2; CE03231.
CC Hypothetical protein.
CC SEQUENCE 495 AA; 55814 MW; 99E9B52032B0B48C CRC64;
SQ
Query Match 60.7%; Score 37; DB 1; Length 495;
Best Local Similarity 45.5%; Pred. No. 23;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
QY 1 PLSIFSRIGD 11
DB 278 PVSGLFKQVGD 288

```

```

RESULT 14
ID MVIN_SYNY3 STANDARD; PRT; 533 AA.
AC 055179;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Virulence factor mvin homolog.
GN MVIN OR SLR0488.
OS Synecocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.
OX NCBI_TaxId=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96127529; PubMed=8550279;
RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA Sugita M., Tabata S.;
RT *Sequence analysis of the genome of the unicellular cyanobacterium
RT Synecocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 644 to 924 of the genome."/
RL DNA Ref. 21153-166 (1995).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC -1- SIMILARITY: BELONGS TO THE MVIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.iesb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: D64001; BAA10319.1;
DR InterPro; IPR004268; MVIN_1like.
DR Pfam; PF03023; MVIN_1.
KW Transmembrane; Complete proteome.
SQ
FT TRANSMEM 11 31 POTENTIAL.
FT TRANSMEM 46 66 POTENTIAL.
FT TRANSMEM 96 116 POTENTIAL.
FT TRANSMEM 135 155 POTENTIAL.
FT TRANSMEM 166 186 POTENTIAL.
FT TRANSMEM 196 216 POTENTIAL.
FT TRANSMEM 253 273 POTENTIAL.
FT TRANSMEM 284 304 POTENTIAL.
FT TRANSMEM 330 350 POTENTIAL.
FT TRANSMEM 360 380 POTENTIAL.
FT TRANSMEM 400 420 POTENTIAL.
FT TRANSMEM 422 442 POTENTIAL.
FT TRANSMEM 452 472 POTENTIAL.
FT TRANSMEM 493 513 POTENTIAL.
SQ SEQUENCE 533 AA; 57053 MW; 00A5B612F744C1D2 CRC64;

Query Match 60.7%; Score 37; DB 1; Length 533;
Best Local Similarity 50.0%; Pred. No. 25;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 PLUSIFSRIGDP 12
Db 300 PFLPVFSRIADP 311

RESULT 15
ID DCUP_CHLMU STANDARD; PRT; 334 AA.
AC 09P1H7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Uroporphyrinogen decarboxylase (EC 4.1.1.37) (URO-D) (UPD).
GN HEME OR TC0123.
OS Chlamydia muridarum.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxId=83360;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Mopn / Nig9;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gyll S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Uterback T., Berry K., Baas S.,
RA Linher K., Weidman J., Knouri H., Craven B., Bowman C., Dodson R.,
RA Gwin M., Nelson W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Frazer C.M.;
RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
RT pneumoniae AR39."
RL Nucleic Acids Res. 28:1397-1406 (2000).
CC -1- CATALYTIC ACTIVITY: Uroporphyrinogen-III = coproporphyrinogen + 4
CC CO(2).
CC -1- PATHWAY: Porphyrin biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (probable).
CC -1- SIMILARITY: BELONGS TO THE UROPORPHYRINOGEN DECARBOXYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.iesb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AE002279; AAF39001.1; -.
DR HSSP; P06132; URO.
DR TIGR; TC0123; -.
DR InterPro; IPR00257; Uro_decarboxyls.
DR Pfam; PF01208; URO-D; 1.
DR PROSITE; PS00906; UROD_1; 1.
DR PROSITE; PS00907; UROD_2; 1.
KW Lyase; Decarboxylase; Porphyrin biosynthesis; Complete proteome.
SQ SEQUENCE 334 AA; 37380 MW; 199B6A8BAC1F4D2D CRC64;

Query Match 59.0%; Score 36; DB 1; Length 334;
Best Local Similarity 54.5%; Pred. No. 24;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 LSSIFSRIGDP 12
Db 258 LSEIYRQLGDP 268

```

Search completed: February 3, 2003, 09:24:21
 Job time : 1.43902 secs

